What Motivates Effort? Evidence and Expert Forecasts
Online Appendix
Forthcoming, *Review of Economic Studies*

Stefano DellaVigna
UC Berkeley and NBER

Devin Pope
U Chicago and NBER

April 2017

A Online Appendix A - Estimation Appendix

Minimum Distance. Given 3 moments (mean effort for no-pay, 1 cent per 100, and 10 cents per 100) and 3 unknowns (γ, k, and s), the system is just identified and we eventually solve 3 equations with 3 unknowns. That is, we solve numerically the system of

$$0 = \gamma * log(\bar{e_p}) - log(s + p) + log(k) \text{ where } p \in \{0, 0.01, 0.10\}$$

Thanks to monotonicity the system has a single solution with the observed mean efforts.

Given the estimates for $\hat{k}$, $\hat{s}$, and $\hat{\gamma}$ (e.g. as in column 1 of Table 5), we then back out the implied behavioral parameters given the mean-effort or predictions of experts. For example, to get the estimates of altruism α and warm-glow a we solve a system of two equations (the mean effort in the two relevant charity treatments) in two unknowns, α and a (equation 11 in the paper). By design, the model is just identified[1]. This way we solve for the five behavioral parameters in columns 1 and 4 of Table 5 panel B. To get the implied parameters forecasted by the experts (e.g. columns 2 and 5 of Table 5 - panel B), we use the same estimates of $\hat{k}$, $\hat{s}$, and $\hat{\gamma}$ (assuming that the experts could have solved for them) and replace the mean efforts in the relevant treatments (e.g., the two charity treatments) with the predictions of each expert for a given treatment; that yields implied behavioral parameters for each expert, i.e., $(\widetilde{\alpha}_i, \widetilde{a}_i)$ for expert i. In the tables we report the 25%, 50%, and 75% statistics of the implied behavioral parameters across the experts.

We derive confidence intervals for the parameters using a bootstrap procedure. We draw 1,000 samples. In each bootstrap iteration we resample (with replacement) workers from each treatment, form a new sample with the same number of observations as the original, and calculate the mean effort. We then re-estimate the parameters $\hat{k}$, $\hat{\gamma}$, and $\hat{s}$ and the relevant behavioral parameters. In panel A of Table 5 we use the standard deviation of the 1000 point estimates, and for panel B as the confidence intervals we use the 2.5th percentile and the 97.5th percentile of the estimated parameters across the 1,000 iterations.

Non-Linear Least Squares. The NLS estimation allows us to model the heterogeneity in effort e_j across the subjects in the experiment. (The minimum-distance estimation uses only the average effort as input.) To take into account the discontinuous incentives, we assume that the individual chooses output in units of 100 points, and estimate the model using output rounded to the closest 100-point: that is, a score of 2,130 points is recorded as 21 units of 100 points. Scores in the 50-99 range are rounded up, while scores in the 0-49 range are rounded down. For the very first bin, that is, scores of 1-49 points, we round to the midpoint, 25.

This assumption allows us to use the first-order condition for effort and thus the non-linear least squares for estimation. Notice that this estimation strategy, while not making use of the full information, is not mis-specified, as it recognizes incentives as actually set. The alternative is to model the continuous point score using maximum-likelihood, model the bunching at the 100-point score. We opted for the simpler and more transparent non-linear least squares estimate.

Unlike in the minimum-distance, here we estimate jointly the behavioral parameters and the cost-function parameters. In Table 5, for example, in Columns 2 and 4 of panel A we report the estimates using only the 3 benchmark treatments. In Columns 3 and 6 of panel B instead we use all 8 relevant treatments (the three benchmark ones, the two on charity, the gift exchange one, and the two time discounting ones). We do not report the cost parameters again for the estimates in Panel B, but they are almost identical to the estimates obtained in Panel A.

In Column 7 we estimate the behavioral parameters for the experts implied by the NLS estimates with exponential cost function. To do this, we follow a similar procedure as for the

[1] Jointly estimating a system of five equations in five unknowns including also the three benchmark treatments yields identical estimates.

minimum-distance estimation. We take the point estimates for the benchmark parameters $\hat{k}$, $\hat{\gamma}$ and $\hat{s}$ from the NLS estimate in Panel A (Column 4). We then take the forecast by expert i for effort in the relevant behavioral treatments (e.g., the two charity treatments) and we use it to infer the implied behavioral parameters for each expert, i.e., $(\widetilde{\alpha}_i, \widetilde{a}_i)$ for expert i. In the tables we report the 25%, 50%, and 75% statistics of the implied behavioral parameters across the experts.

Loss Aversion. A threshold payment (such as at 2,000 points) induces bunching at the threshold, and a missing mass to the left of the threshold. The extent of bunching and the missing mass will be increasing in the utility gain u of achieving the threshold. Denote with $F(u)$ the share bunching as a function of the utility benefit u which, as derived in Section 2 in the paper, equals $(1+\eta)\,G$ in the gain treatment and $(1+\lambda\eta)\,G$ in the loss treatment. The average effort e in a treatment will be an increasing function ϕ of the bunching, and thus $e = g(u)$, where $g(.) \equiv \phi(F(.))$ is an increasing function. Consider a linear approximation to how the average effort responds to a change in u: $de = g'(u^*) * du$. The effort change going from the 40-cent gain condition to the 80-cent gain condition is approximately: $e_{G.80} - e_{G.40} \simeq g'(u^*)\,[(1+\eta)\,.80 - (1+\eta)\,.40] = g'(u^*)\,(1+\eta) * .40$. Similarly, the effort change going from the 40-cent gain condition to the 40-cent loss condition is approximated as $e_{L.40} - e_{G.40} \simeq g'(u^*)\,[(1+\lambda\eta)\,.40 - (1+\eta)\,.40] = g'(u^*)\,(\lambda-1)\,\eta * .40$. The ratio of these differences is

$$\frac{e_{L.40} - e_{G.40}}{e_{G.80} - e_{G.40}} \simeq \frac{g'(u^*)\,(\lambda-1)\,\eta * .40}{g'(u^*)\,(1+\eta) * .40} = \frac{(\lambda-1)\,\eta}{1+\eta}.$$

The term $g'(u^*)$ drops out, leaving a function of just λ and η. Under the standard assumption of unitary gain utility ($\eta = 1$), the ratio of the difference in effort allows for estimation of the loss aversion λ. Notice that, unlike the other derivations, this solution is an approximation. However, given that the differences in effort between the threshold treatments are small, the bias in estimate due to the approximation should be small as well.

Given that the estimation is based on a ratio, we only use observations in which the denominator is positive and larger than 10 units of effort, since smaller differences may be hard for experts to even control with a mouse, and we exclude observations with negative λ.

Low-Pay Treatment. The predicted effort for the low-pay treatment in Table 5 assumes for simplicity that the incentive (1-cent every 1,000 points) is paid continuously, as opposed to only at every 1,000-point threshold. This is true also for the NLS specification, which assumes an incentive of 0.1 cents every 100 points in order to be apply the first order conditions.

We now show that modelling the payoff jumps at the 1,000-point thresholds leads to similar predicted effort for the low-pay condition. Consider the NLS estimate with exponential cost of effort function. (Modelling the threshold effects only makes sense for models with heterogeneity, that is, the NLS models and not the minimum-distance model). Individual i maximizes

$$\max_{e_i} se_i + p(\mathbf{1}_{e_i \geq 1000} + \mathbf{1}_{e_i \geq 2000} + \mathbf{1}_{e_i \geq 3000}) - c_i(e_i)$$

where $c_i(e_i)$ is specific to person i: $c_i(e_i) = k \exp(\gamma e_i)\,\gamma^{-1} \exp(-\gamma \varepsilon_i) = (k/\gamma) \exp(\gamma (e_i - \varepsilon_i))$. The second term models the threshold compensation p (which equals 1 cent in this case). For expositional simplicity, we assume that exceeding 4,000 points is too costly. Consider the optimal solution without incentives ($p = 0$): $s - k \exp(\gamma e_i) \exp(-\gamma \varepsilon_i) = 0$ or

$$e_i = \frac{1}{\gamma}\,[\log(s/k)] + \varepsilon_i. \tag{1}$$

Given the estimated $\hat{s}$, $\hat{\gamma}$, and $\hat{k}$, the realization of ε_i pins down uniquely e_i. Thus, denote with $\varepsilon_i(e_i)$ the error term that leads to the choice of e_i with no incentives. We now show that the solution takes a threshold form, which we characterize first with respect to the threshold at 1,000 points. There is a value ε^{1000} such that any type $\varepsilon < \varepsilon^{1000}$ stays at the effort level chosen with no incentive as in (1). Any type with $\varepsilon > \varepsilon^{1000}$ chooses instead to jump to the threshold effort of 1,000 or stay at his already higher level of effort. Label as $U_0(e_i)$ the utility that type e_i achieves at the optimum as in (1), substituting for the expression for $\varepsilon(e_i)$ and simplifying, we obtain

$$U_0(e_i) = se_i - \frac{k}{\gamma}\exp(\gamma e_i)\exp\left(-\gamma e_i + \log\left(\frac{s}{k}\right)\right) = s\left(\bar{e} - \frac{1}{\gamma}\right).$$

Label as U_{1000} the utility that type e_i achieves from exerting effort 1,000. With similar substitutions and simplifications we obtain

$$U_{1000}(e_i) = 1000s + 1 - \frac{s}{\gamma}\exp(\gamma(1000 - \bar{e})).$$

We show now that there exists one and only one $\bar{e}$, with $\bar{e} < 1000$, such that $U_0(\bar{e}) = U_{1000}(\bar{e})$. Furthermore, $U_0(e_i) > U_{1000}(e_i)$ for $e_i < \bar{e}$ and $U_0(e_i) < U_{1000}(e_i)$ for $1000 > e_i > \bar{e}$, that is, there is a threshold strategy.

First, note that $U_0(1000) = s(1000 - 1/\gamma) < U_{1000}(1000) = s(1000 - 1/\gamma) + 1$. Thus by continuity, types e_i close enough to 1000 will strictly prefer 1000 to the solution in (1). Then notice that $U_0(e_i)$ increases linearly in e_i with derivative s, while the derivative of U_{1000} with respect to e_i is

$$\frac{\partial U_{1000}(e_i)}{\partial e_i} = s\exp(\gamma(1000 - e_i)) > 0.$$

This derivative is decreasing in e_i, that is, U_{1000} is an increasing and concave function of e_i. Furthermore, for negative enough e_i the derivative $\partial U_{1000}(e_i)/\partial e_i$ becomes arbitrarily large and thus larger than the derivative $\partial U_0(e_i)/\partial e_i$. Thus, for e_i small enough, it must be the case that $U_0(e_i) > U_{1000}(e_i)$. To show that there is only one point of crossing between U_0 and U_{1000} for $e < 1000$, consider once again the properties of the two derivative functions. This concludes the proof.

Having determined the threshold $\bar{e}^{1000}$ we can similarly derive the other thresholds $\bar{e}^{2000}$ and $\bar{e}^{3000}$. Thus we know that the observed distribution will consist of a mixture of density from 0 to $\bar{e}^{1000}$, bunching at 1,000, then density from 1,000 to $\bar{e}^{2000}$ and so on. For the estimated $\hat{s}, \hat{\gamma}$, and $\hat{k}$, the threshold expressed in effort units are 185 (so types with effort higher than 185 and lower than 1,000 will jump to 1,000), 1,130, and 2,097.

The only remaining piece to determine is the distribution of the error term ε_i. We present the results following two approaches. The first approach just takes the estimated standard deviation of ε from the non-linear least squares estimation. The second approach instead backs out the distribution of ε non-parametrically from the no-payment case: an observed e_i, given the estimated $\hat{s}, \hat{\gamma}$, and $\hat{k}$, implies a realization of ε_i. Under either approach, we compute the counterfactual effort for the low-pay treatment, by moving the observations which are predicted to bunch to the bunching point, and then compute the expected effort. The first approach yields a simulated mean effort of 1,881, while the second approach yields a similar counterfactual of 1,878 for the mean effort. Thus, the effort is similar to the counterfactual estimated assuming continuous point earning.

Meta-Analysis. This subsection presents additional details on computations related to the meta-analysis.

First, we walk through the calculation of Cohen's d. Using statistics reported in most papers, we typically recover the means, variances and numbers of subjects in the treatment and control groups. When the study design uses random assignment (which is typically the case in the papers we found), the treatment effect is identified as the difference in group means. In order to compare this treatment effect with those found in other papers, we normalize it into a unitless quantity, dividing by the pooled standard deviation (which is the square root of a weighted sum of variances for the control and treatment groups, where the weights are proportional to the number of observations in the group up to a degrees-of-freedom correction). This calculated value is known as Cohen's d – a measure of effect size that is commonly used in meta-analyses. There is also a standard error associated with Cohen's d (given by a standard formula), which is decreasing in the numbers of observations in the control and treatment groups, and increasing in the magnitude of Cohen's d.

The procedure described above yields an estimate and standard error for each paper's Cohen's d. In order to aggregate these estimates into a single value for each treatment (e.g. gift exchange), we use two alternative approaches. Our preferred approach (reported in column 8 of table 4) is to use the inverse-variance weighted mean of the Cohen's d's from papers associated with the treatment (where the variances are the squares of the Cohen's d's standard errors). This results in greater weights for more precisely estimated effect sizes or similarly, for studies with larger sample sizes. As robustness, we also tried a citation-weighted mean of Cohen's d's (reported in column 9 of table 4), where we use Google Scholar citations as of late 2016. Google scholar citations are preferable to citations from the Social Science Citation Index for our purposes because we include in our meta-analysis working papers that have not yet been published. The citation-weighted mean gives more weight to papers that are better known in the literature (as proxied by citations).

The aggregated Cohen's d for each treatment can be compared to our MTurk experimental results and expert forecasts in a couple of ways:

First, we can convert MTurks' effort or experts' forecasts into units of Cohen's d and compare them to the effect size implied by the literature. We use the "gift exchange versus no piece rate" comparison as an example to illustrate this procedure. Our control group consists of MTurks who were offered no piece rate, whereas the treatment group are MTurks who were unconditionally given 40 cents. The calculation of the Cohen's d implied by MTurks' effort is completely analogous to the description for other papers in the literature. To compute the Cohen's d implied by experts' forecasts, we simply replace the treatment group mean from the MTurk calculation with the average expert forecast for the gift exchange treatment, then follow same calculations above. These results are reported in columns 3 and 4 of table 4.

Second, we can convert the meta-analysis effect sizes into forecasts about implied effort by MTurks for various treatment comparisons if they had behaved in line with the literature. Returning to the "gift exchange versus no piece rate" example, we take the MTurk effort in the "no piece rate" treatment and the pooled standard deviation for the control and treatment groups as given. The literature-implied effort for MTurks in the "gift exchange" group is then the sum of the mean control group effort and the product of the literature implied Cohen's d with the pooled standard deviation. These are the computations underlying figure 8 and online appendix figure 7 in the paper.

References

[1] Abdellaoui, Mohammed. 2000. "Parameter-free elicitation of utility and probability weighting functions." Management Science 46(11): 1497-1512.

[2] Ariely, Dan, Emir Kamenica, and Dražen Prelec. 2008. "Man's search for meaning: The case of Legos." Journal of Economic Behavior & Organization 67(3): 671-677.

[3] Aruoba, S. Boragan, and Melissa Schettini Kearney. 2010. Do lottery gamblers love risk or overweight small odds? University of Maryland working paper.

[4] Ashraf, Nava, Bandiera, Oriana and Jack, B. Kelsey. 2014. "No margin, no mission? A field experiment on incentives for public service delivery." Journal of Public Economics, 120: 1-17.

[5] Ashraf, Nava, Oriana Bandiera, and Scott S. Lee. 2014. "Awards unbundled: Evidence from a natural field experiment." Journal of Economic Behavior & Organization 100: 44-63.

[6] Barankay, Iwan. 2012. "Rank incentives: Evidence from a randomized workplace experiment." Working paper.

[7] Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue, and Joshua C. Teitelbaum. 2013. "The nature of risk preferences: Evidence from insurance choices." The American Economic Review 103(6): 2499-2529.

[8] Bellemare, Charles, and Bruce Shearer. 2011. "On the relevance and composition of gifts within the firm: Evidence from field experiments." International Economic Review 52(3): 855-882.

[9] Beshears, John, James J. Choi, David Laibson, Brigitte C. Madrian, and Katherine L. Milkman. 2015. "The effect of providing peer information on retirement savings decisions." The Journal of finance 70(3): 1161-1201.

[10] Bhargava, Saurabh, and Dayanand Manoli. 2015. "Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment." The American Economic Review 105(11): 3489-3529.

[11] Blanes i Vidal, Jordi, and Mareike Nossol. 2011. "Tournaments without prizes: Evidence from personnel records." Management Science 57(10): 1721-1736.

[12] Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper. 2010. "Risk and rationality: Uncovering heterogeneity in probability distortion." Econometrica 78(4): 1375-1412.

[13] Cai, Hongbin, Yuyu Chen, and Hanming Fang. 2009. "Observational learning: Evidence from a randomized natural field experiment." The American Economic Review 99(3): 864-882.

[14] Camerer, Colin F., and Teck-Hua Ho. 1994. "Violations of the betweenness axiom and nonlinearity in probability." Journal of Risk and Uncertainty 8(2): 167-196.

[15] Chandler, Dana, and Adam Kapelner. 2013. "Breaking monotony with meaning: Motivation in crowdsourcing markets." Journal of Economic Behavior & Organization 90: 123-133.

[16] Charness, Gary, Ramón Cobo-Reyes, and Ángela Sánchez. 2016. “The effect of charitable giving on workers’ performance: Experimental evidence,” Journal of Economic Behavior and Organization, 131, 61-74.

[17] Coffman, Lucas C., Clayton R. Featherstone, and Judd B. Kessler. 2017. ”Can Social Information Affect What Job You Choose and Keep?.” *AEJ: Applied Economics*, Vol. 9, pp. 96-117.

[18] Cohn, Alain, Ernst Fehr, and Lorenz Goette. 2014. ”Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment.” Management Science 61(8): 1777-1794.

[19] De Brauw, Alan, and Patrick Eozenou. 2014. ”Measuring risk attitudes among Mozambican farmers.” Journal of Development Economics 111: 61-74.

[20] Deehan, Ann, Lorna Templeton, Colin Taylor, Colin Drummond, and John Strang. 1997. ”The effect of cash and other financial inducements on the response rate of general practitioners in a national postal study.” British Journal of General Practice 47(415): 87-90.

[21] DellaVigna, S., List, J.A., Malmendier, U. and Rao, G., 2016. “Estimating Social Preferences and Gift Exchange at Work.” National Bureau of Economic Research Working paper 22043.

[22] Diamond, William D., and Ben Z. Loewy. 1991. ”Effects of Probabilistic Rewards on Recycling Attitudes and Behavior.” Journal of Applied Social Psychology 21(19): 1590-1607.

[23] Dolan, Paul, and Caroline Rudisill. 2014. ”The effect of financial incentives on Chlamydia testing rates: evidence from a randomized experiment.” Social Science & Medicine 105: 140-148.

[24] Donkers, Bas, Bertrand Melenberg, and Arthur Van Soest. 2001. ”Estimating risk attitudes using lotteries: A large sample approach.” Journal of Risk and uncertainty 22(2): 165-195.

[25] Englmaier, Florian and Leider, Stephen. 2010. “Gift Exchange in the Lab - It is Not (Only) How Much You Give...” CESifo Working Paper Series No. 2944.

[26] Englmaier, Florian, and Stephen Leider. 2012. ”Managerial payoff and gift exchange in the field.” Working paper.

[27] Esteves-Sorenson, Constanca. 2015. “Gift exchange in the workplace: Addressing the conflicting evidence with a careful test.” Working paper.

[28] Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler. 2013. ”Testing enforcement strategies in the field: Threat, moral appeal and social information.” Journal of the European Economic Association 11(3): 634-660.

[29] Frey, Bruno S., and Stephan Meier. 2004. ”Social comparisons and pro-social behavior: Testing” conditional cooperation” in a field experiment.” The American Economic Review 94(5): 1717-1722.

[30] Gilchrist, Duncan S., Michael Luca, and Deepak Malhotra. 2016. ”When 3+ 1> 4: Gift structure and reciprocity in the field.” Management Science 62(9):2639-2650.

[31] Gill, David, Zdenka Kissová, Jaesun Lee, and Victoria L. Prowse. 2016. "First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision." Working paper.

[32] Gneezy, Uri, and John A. List. 2006. "Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments." Econometrica 74(5): 1365-1384.

[33] Gneezy, Uri, and Pedro Rey-Biel. 2014. "On the relative efficiency of performance pay and noncontingent incentives." Journal of the European Economic Association 12(1): 62-72.

[34] Gneezy, Uri, and Rustichini, Aldo. 2000. "Pay Enough or Don't Pay at All." The Quarterly Journal of Economics, 115(3): 791-810.

[35] Goldstein, Noah J., Robert B. Cialdini, and Vladas Griskevicius. 2008. "A room with a viewpoint: Using social norms to motivate environmental conservation in hotels." Journal of Consumer Research 35(3): 472-482.

[36] Gonzalez, Richard, and George Wu. 1999. "On the shape of the probability weighting function." Cognitive Psychology 38(1): 129-166.

[37] Grant, Adam M. 2008. "The significance of task significance: Job performance effects, relational mechanisms, and boundary conditions." Journal of Applied Psychology 93(1): 108.

[38] Grant, Adam M. 2012. "Leading with meaning: Beneficiary contact, prosocial impact, and the performance effects of transformational leadership." Academy of Management Journal 55(2): 458-476.

[39] Grant, Adam M., Elizabeth M. Campbell, Grace Chen, Keenan Cottone, David Lapedis, and Karen Lee. 2007. "Impact and the art of motivation maintenance: The effects of contact with beneficiaries on persistence behavior." Organizational Behavior and Human Decision Processes 103(1): 53-67.

[40] Hallsworth, Michael, John List, Robert Metcalfe, and Ivo Vlaev. 2014. "The behavioralist as tax collector: Using natural field experiments to enhance tax compliance." National Bureau of Economic Research Working paper 20007.

[41] Halpern, Scott D., Rachel Kohn, Aaron Dornbrand-Lo, Thomas Metkus, David A. Asch, and Kevin G. Volpp. 2011. "Lottery-Based versus Fixed Incentives to Increase Clinicians' Response to Surveys." Health services research 46(5): 1663-1674.

[42] Harrison, Glenn W., John A. List, and Charles Towe. 2007. "Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion." Econometrica 75(2): 433-458.

[43] Harrison, Glenn W., Steven J. Humphrey, and Arjan Verschoor. 2010. "Choice under uncertainty: evidence from Ethiopia, India and Uganda." The Economic Journal 120(543): 80-104.

[44] Hossain, Tanjim and Li, King King. 2013. Crowding out in the labor market: a prosocial setting is necessary. Management Science, 60(5): 1148-1160.

[45] Imas, Alex. 2014. "Working for the "warm glow": On the benefits and limits of prosocial incentives." Journal of Public Economics, 114: 14-18.

[46] Kilka, Michael, and Martin Weber. 2001. "What determines the shape of the probability weighting function under uncertainty?." Management Science 47(12): 1712-1726.

[47] Kliger, Doron, and Ori Levy. 2009. "Theories of choice under risk: Insights from financial markets." Journal of Economic Behavior & Organization 71(2): 330-346.

[48] Liu, Elaine M. 2013. "Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China." Review of Economics and Statistics 95(4): 1386-1403.

[49] Kosfeld, Michael, and Susanne Neckermann. 2011. "Getting more work for nothing? Symbolic awards and worker performance." American Economic Journal: Microeconomics 3(3): 86-99.

[50] Krupka, Erin, and Roberto A. Weber. 2009. "The focusing and informational effects of norms on pro-social behavior." Journal of Economic psychology 30(3): 307-320.

[51] Kube, Sebastian, Michel André Maréchal, and Clemens Puppe. 2013. "Do wage cuts damage work morale? Evidence from a natural field experiment." Journal of the European Economic Association 11(4): 853-870.

[52] Kube, Sebastian, Michel André Maréchal, and Clemens Puppea. 2012. "The currency of reciprocity: Gift exchange in the workplace." The American Economic Review 102(4): 1644-1662.

[53] Snowberg, Erik, and Justin Wolfers. 2010. "Explaining the Favorite–Long Shot Bias: Is it Risk-Love or Misperceptions?." Journal of Political Economy 118(4): 723-746.

[54] Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen. 2006. "Risk and time preferences: linking experimental and household survey data from Vietnam." In Behavioral Economics of Preferences, Choices, and Happiness, edited by Shinsuke Ikeda, Hideaki Kiyoshi Kato, Fumio Ohtake and Yoshiro Tsutsui, 3-25. Japan: Springer.

[55] Thirumurthy, Harsha, Samuel H. Masters, Samwel Rao, Kate Murray, Ram Prasad, Joshua G. Zivin, Eunice Omanga, and Kawango Agot. 2016. "The effects of providing fixed compensation and lottery-based rewards on uptake of medical male circumcision in Kenya: a randomized trial." Journal of Acquired Immune Deficiency Syndromes 72(Supplement 4): S309.

[56] Tonin, Mirco and Vlassopoulos, Michael. 2014. "Corporate philanthropy and productivity: Evidence from an online real effort experiment." Management Science, 61(8): 1795-1811.

[57] Tversky, Amos, and Craig R. Fox. 1995. "Weighing risk and uncertainty." Psychological review 102(2): 269-283.

[58] Tversky, Amos, and Daniel Kahneman. 1992. "Advances in prospect theory: Cumulative representation of uncertainty." Journal of Risk and Uncertainty 5(4): 297-323.

[59] Wu, George, and Richard Gonzalez. 1996. "Curvature of the probability weighting function." Management Science 42(12): 1676-1690.

[60] Yang, Adelle Xue, Christopher K. Hsee, and Oleg Urminsky. 2014. "Eager to help yet reluctant to give: How pro-social effort and pro-social choices diverge." Working paper.

Online Appendix Figures 1a-d. MTurk Task, Examples of Screenshots

Online Appendix Figure 1a. Screenshot for 10-cent benchmark treatment, Instructions

On the next page you will play a simple button-pressing task. The object of this task is to alternately press the 'a' and 'b' buttons on your keyboard as quickly as possible for 10 minutes. Every time you successfully press the 'a' and then the 'b' button, you will receive a point. Note that points will only be rewarded when you **alternate** button pushes: just pressing the 'a' or 'b' button without alternating between the two will not result in points.

Buttons must be pressed by hand only (key-bindings or automated button-pushing programs/scripts cannot be used) or task will not be approved.

Feel free to score as many points as you can.

As a bonus, you will be paid an extra 10 cents for every 100 points that you score. This bonus will be paid to your account within 24 hours.

Online Appendix Figure 1b. Screenshot for 10-cent benchmark treatment, Task

08 57

Press 'a' then 'b'...

Points: 302
Bonus Payout: $ 0.30

You will be paid an extra 10 cents for every 100 points that you score.

Online Appendix Figure 1c. Screenshot for 40-cent gain treatment, Instructions

On the next page you will play a simple button-pressing task. The object of this task is to alternately press the 'a' and 'b' buttons on your keyboard as quickly as possible for 10 minutes. Every time you successfully press the 'a' and then the 'b' button, you will receive a point. Note that points will only be rewarded when you **alternate** button pushes: just pressing the 'a' or 'b' button without alternating between the two will not result in points.

Buttons must be pressed by hand only (key-bindings or automated button-pushing programs/scripts cannot be used) or task will not be approved.

Feel free to score as many points as you can.

As a bonus, you will be paid an extra 40 cents if you score at least 2,000 points. This bonus will be paid to your account within 24 hours.

Online Appendix Figure 1d. Screenshot for 40-cent gain treatment, Task

09 17

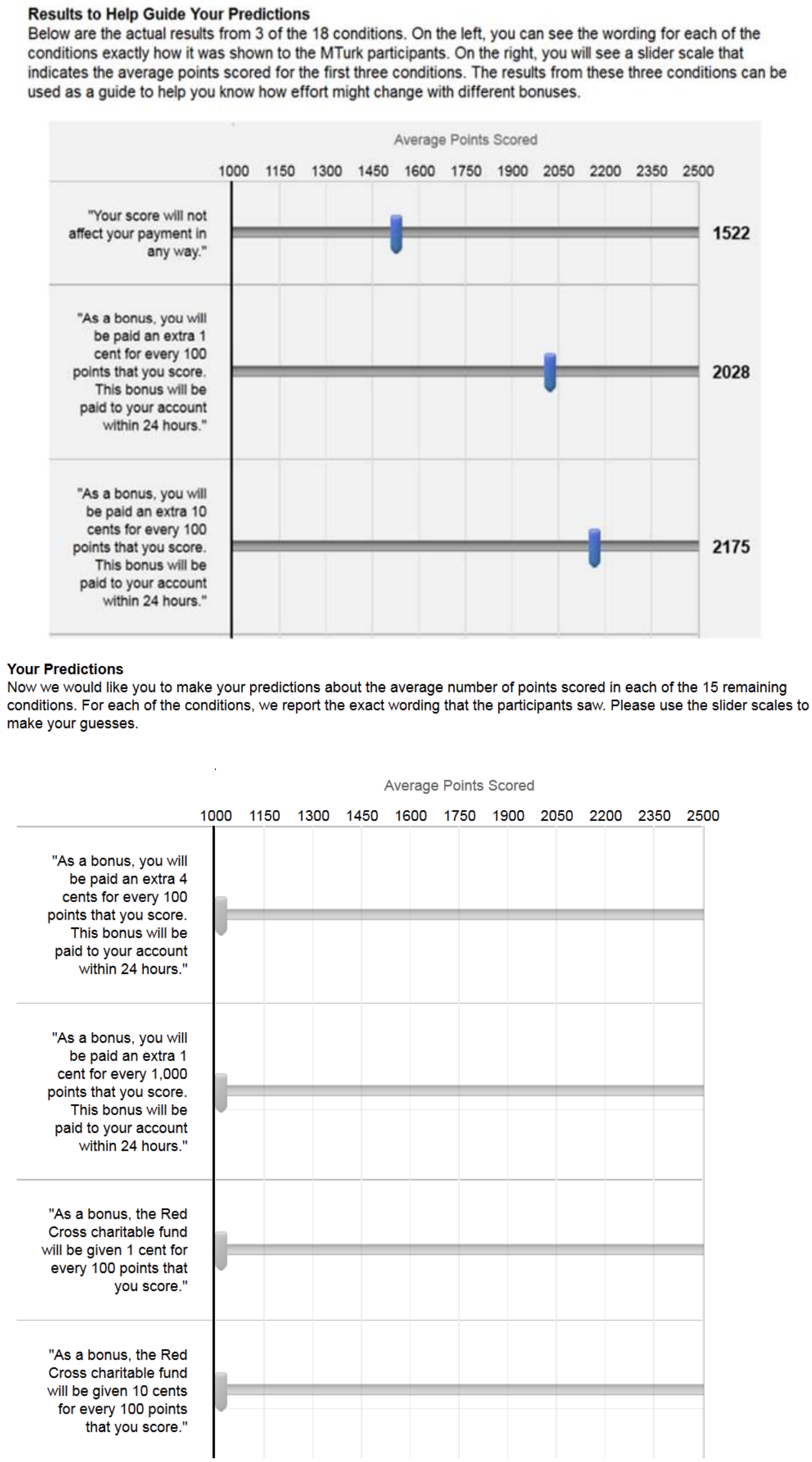
Press 'a' then 'b'...

Points: 215
Bonus Payout: $ 0.00

You will be paid an extra 40 cents if you score at least 2,000 points.

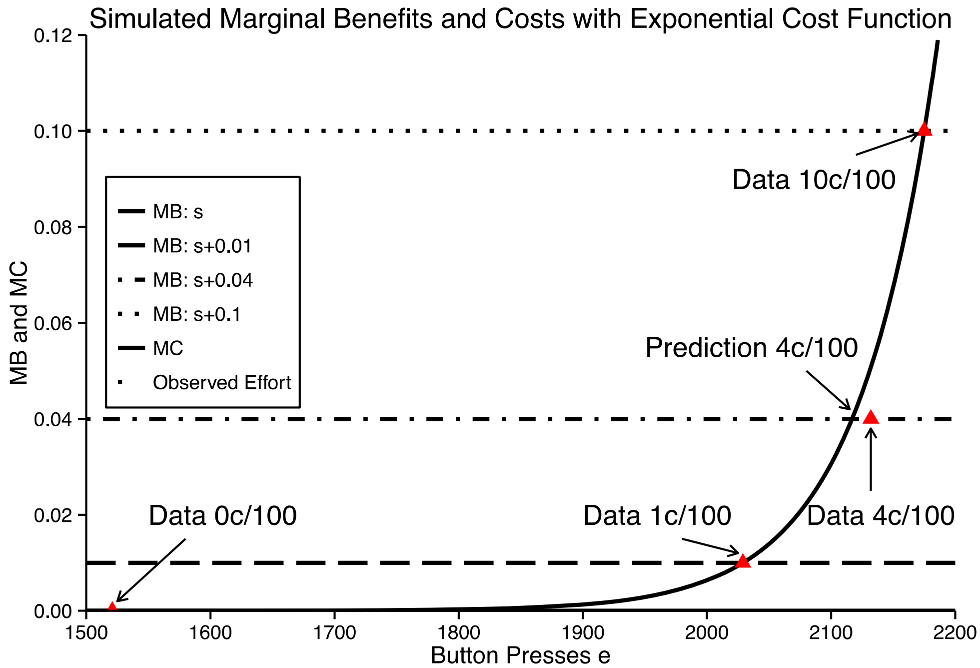
Notes: Online Appendix Figures 1a-d plot excerpts of the MTurk real-effort task for two treatments, the 10-cent piece rate benchmark treatment (Appendix Figure 1a-b) and the 40-cent gain treatment (Appendix Figure 1c-d). For each treatment, the first screenshot reproduces partially the instructions, while the second screenshot displays the task. These two screens are the only places in which the treatments differed.

Online Appendix Figure 2. Expert Survey, Screenshot

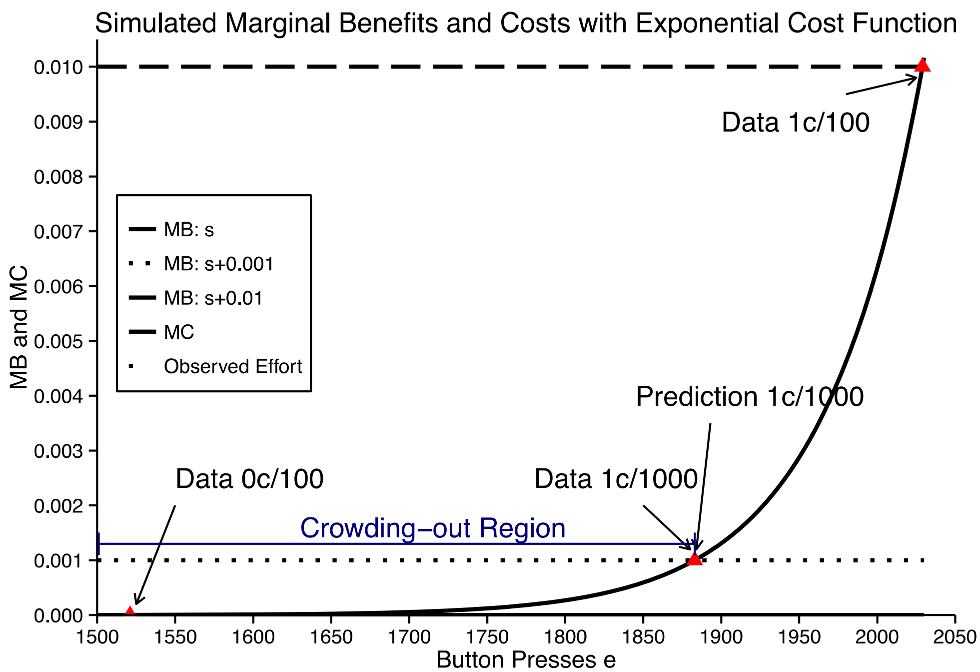


Notes: Online Appendix Figure 2 shows two screenshots reproducing portions of the Qualtrics survey which experts used to make forecasts. The first screenshot reproduces the information provided to the experts about the 3 benchmark treatments. The second screenshot shows 3 of 15 sliders, one for each treatment. For each treatment, the left side displays the treatment-specific wording which the subjects assigned to that treatment saw, and on the right side a slider which the experts can move to make a forecast.

Online Appendix Figure 3. Estimate of Model, Alternative Cost Function (Exponential Cost Function)
Online Appendix Figure 3a. Estimate with 0c, 1c, 10c Piece Rate, Prediction for 4c Piece Rate (Exponential)

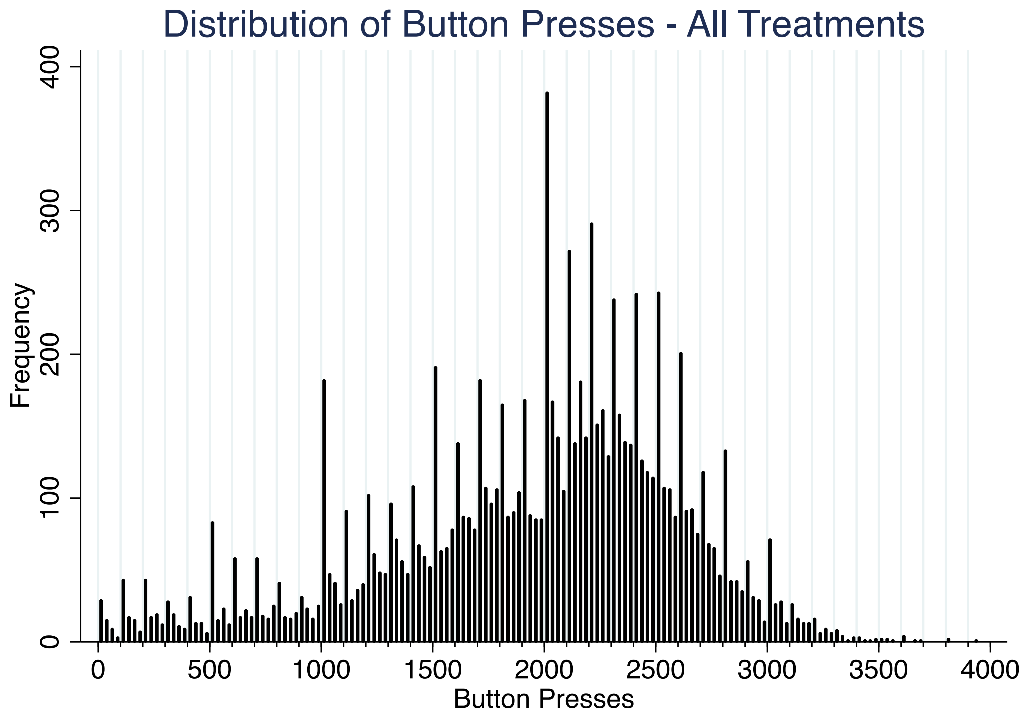


Online Appendix Figure 3b. Predicted Effort for "Paying Too Little" treatment (Exponential)



Notes: Online Appendix Figures 3a-b plot the equivalent of Figures 2a-b, but estimated with an exponential cost function as opposed to a power cost function. Online Appendix Figure 3a plots the marginal cost curve and the marginal benefit curve for the three benchmark treatments. The figure also plots the out of sample prediction for the 4 cent treatment (which is not used in the estimates), as well as the observed effort for that treatment. Online Appendix Figure 3b plots, for the same point estimates, the out of sample prediction for the treatment with 1-cent per 1,000 clicks.

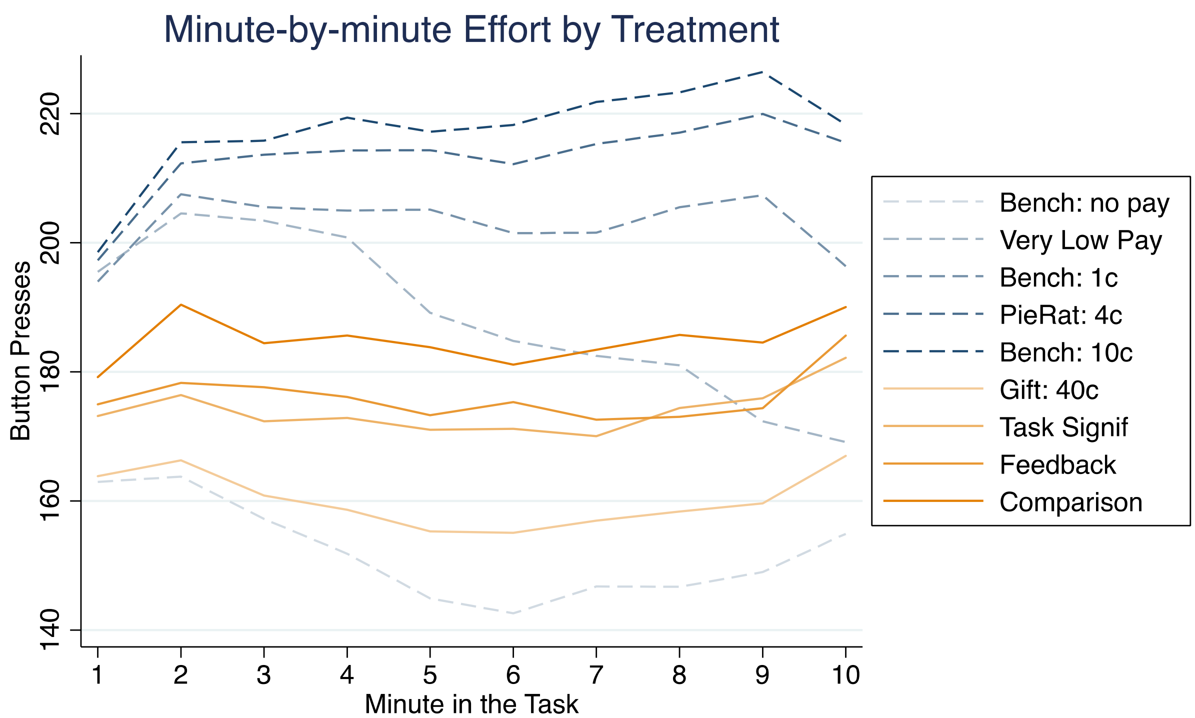
Online Appendix Figure 4. Distribution of Button Presses, All Treatments



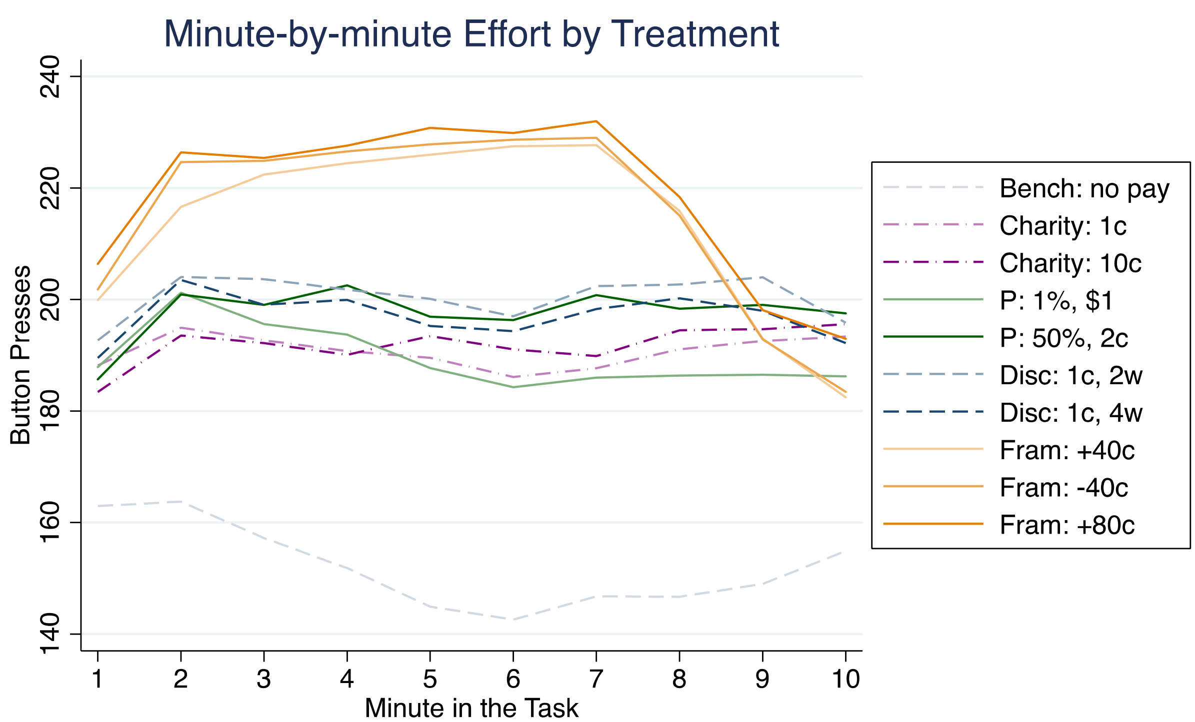
Notes: Online Appendix Figure 4 plots a histogram of the observed button presses over all 18 treatments in the real-effort MTurk experiment in bins of 25 points. Notice the spikes at round numbers, in part because incentives kick in at round-number points.

Online Appendix Figure 5. Effort over Time, MTurk Workers

Online Appendix Figure 5a. Treatments with no Incentives and Piece Rate Treatments

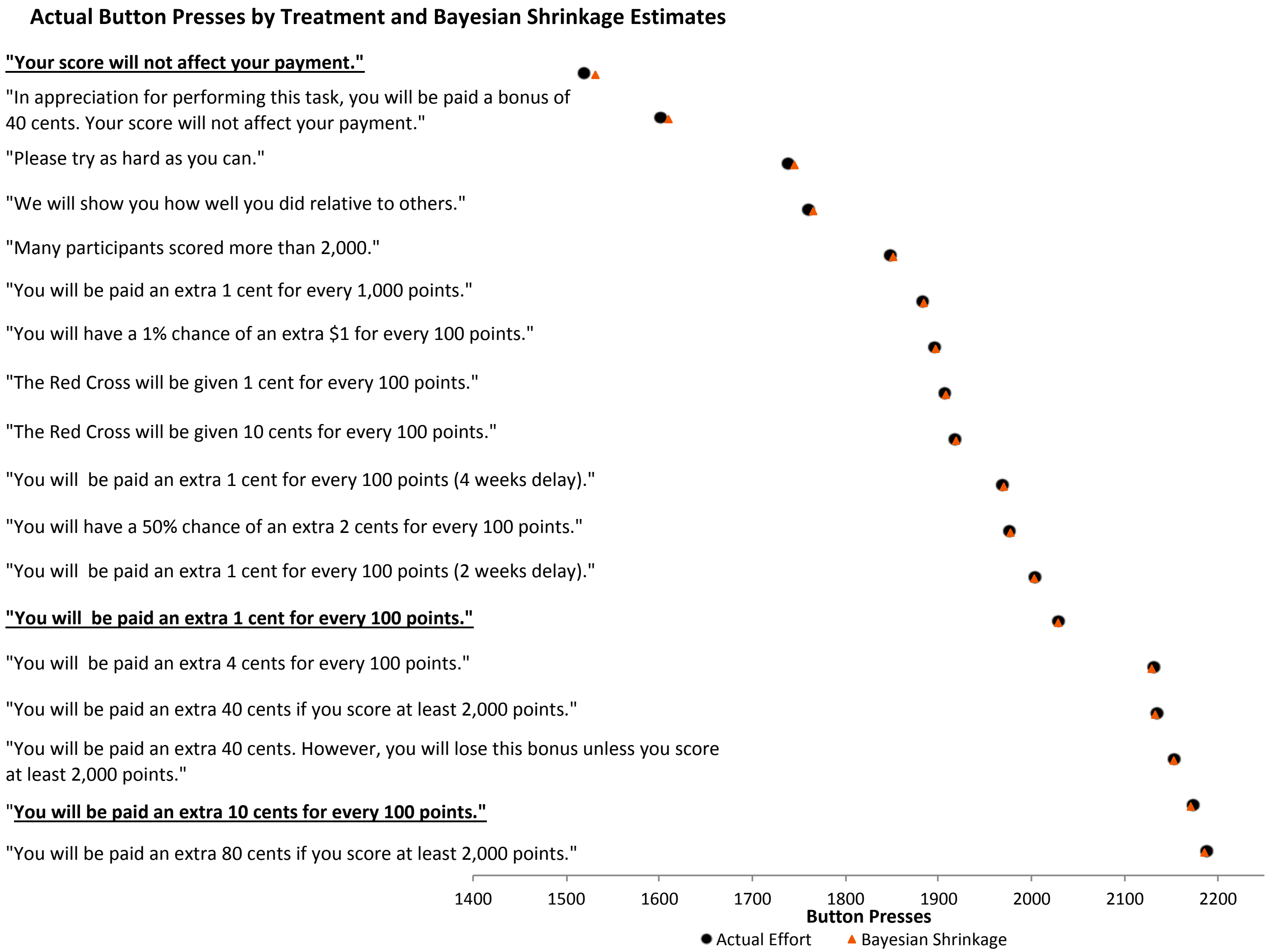


Online Appendix Figure 5b. Other Treatments



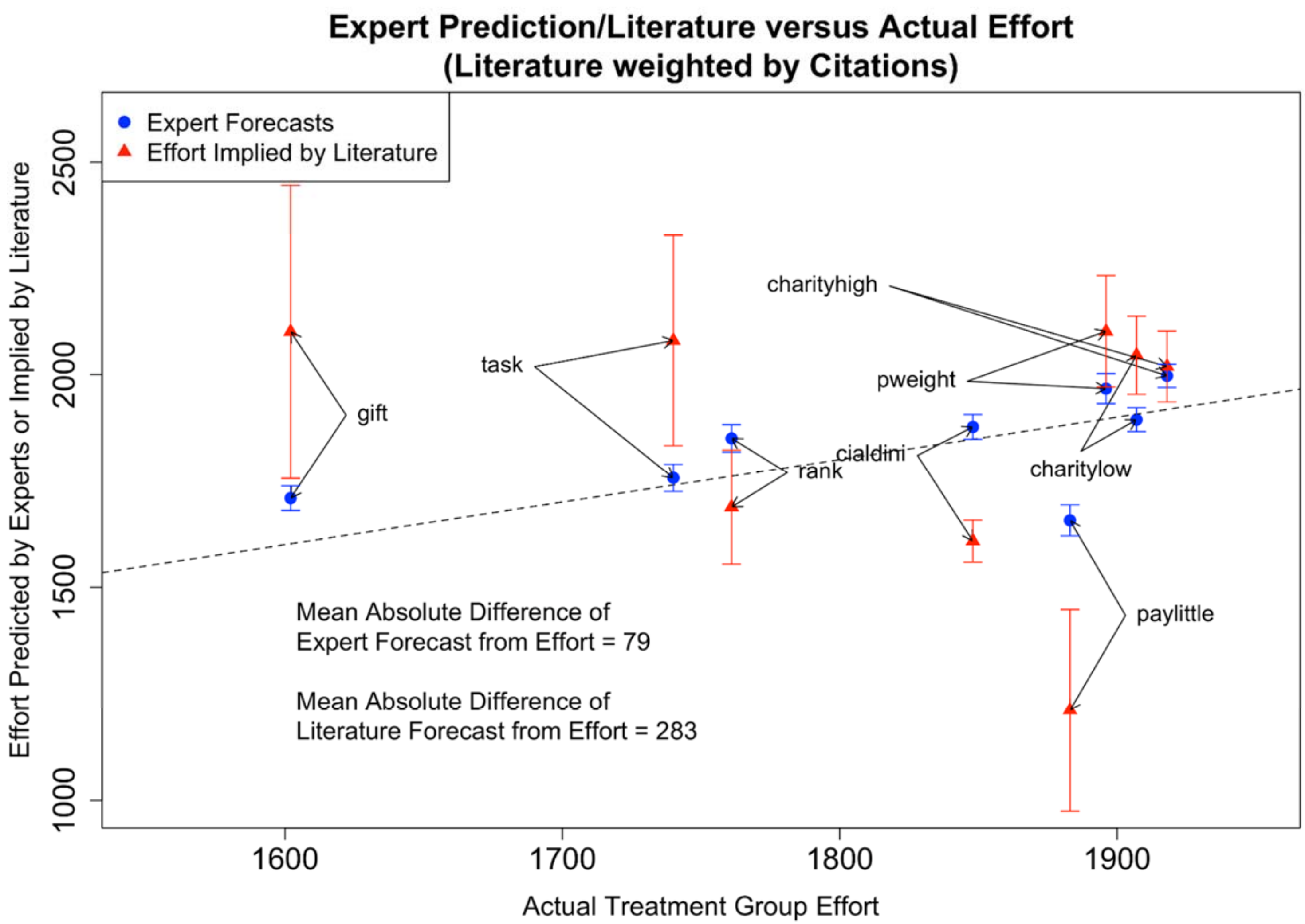
Notes: Online Appendix Figure 5 presents the effort over time for selected treatments. The y axis indicates the average number of button presses in that treatment per minute.

Online Appendix Figure 6. Effort by Treatment, Average and Bayesian Shrinkage Estimator



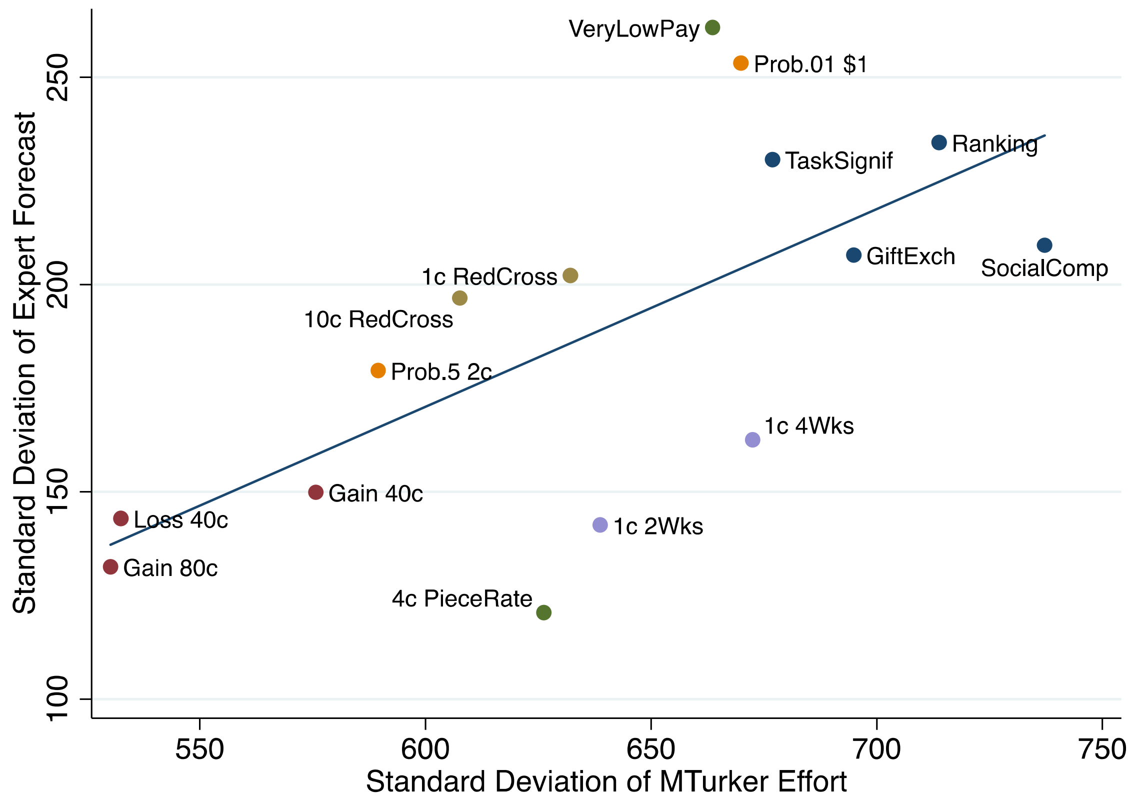
Notes: Online Appendix Figure 6 plots the average effort by treatment as in Figure 3, with in addition a Bayesian shrinkage-adjusted measure, to correct for the sampling error (see text for detail). The adjustment makes only a minimal difference.

Online Appendix Figure 7. Prediction based on literature meta-analysis vs. Expert Forecasts, Citation-based Weights



Notes: Online Appendix Figure 7 presents a parallel to Figure 8 in the text, except that the effort implied by the literature on the y axis is computed using the citation-weighted Cohen's d, instead of the inverse variance-weighted Cohen's d. This citation-weighted meta-analysis makes noisier, and more incorrect, predictions.

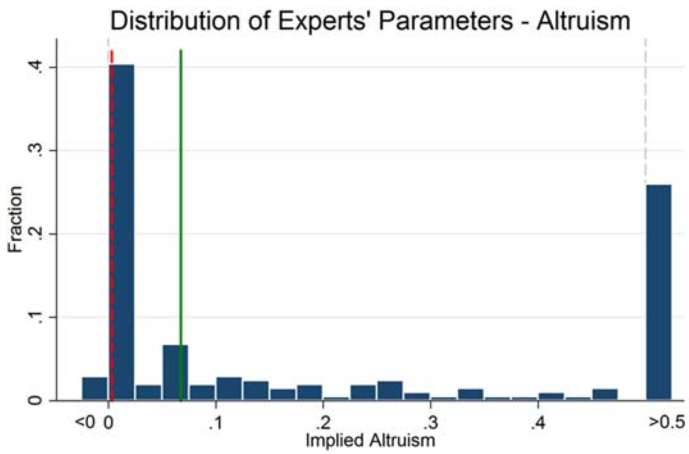
Online Appendix Figure 8. Heterogeneity of Expert Forecasts and Heterogeneity of MTurker Effort, by Treatment



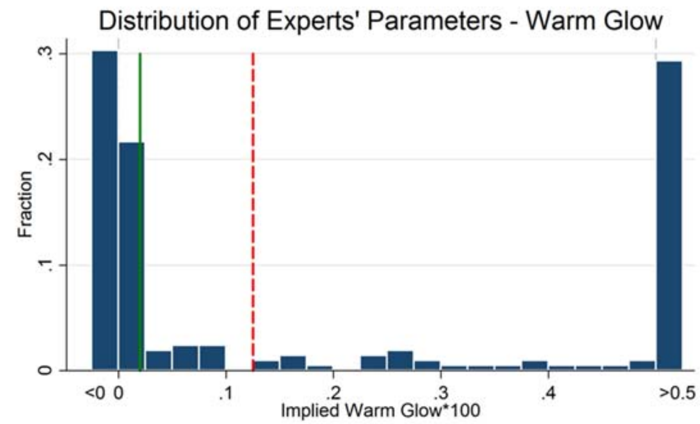
Notes: Online Appendix Figure 8 presents a scatterplot of the 15 treatments, with the standard deviation in MTurker effort on the x axis and the standard deviation in the expert forecast on the y axis. The figure also displays the best-fit line.

Online Appendix Figure 9. Structural Estimates of Behavioral Parameters: Data versus Experts Beliefs

Online Appendix Figures 9a-b. Estimate of Social Preference Parameters

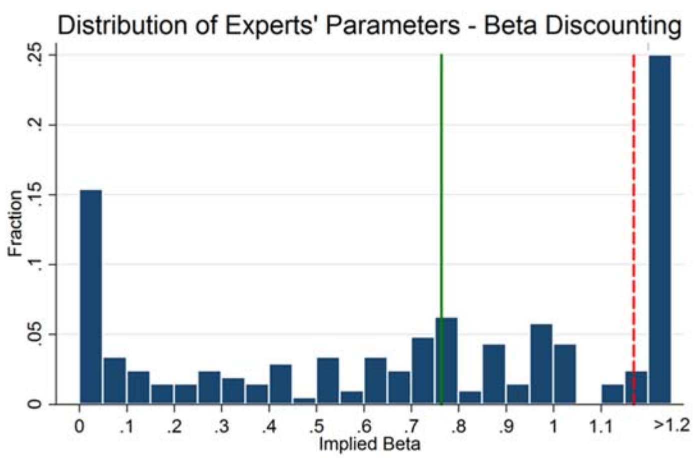


Red dashed line indicates the parameter value implied by the actual mTurker mean effort.
Green solid line indicates the median parameter value of the experts.
Observations are truncated at 0 and 0.5.

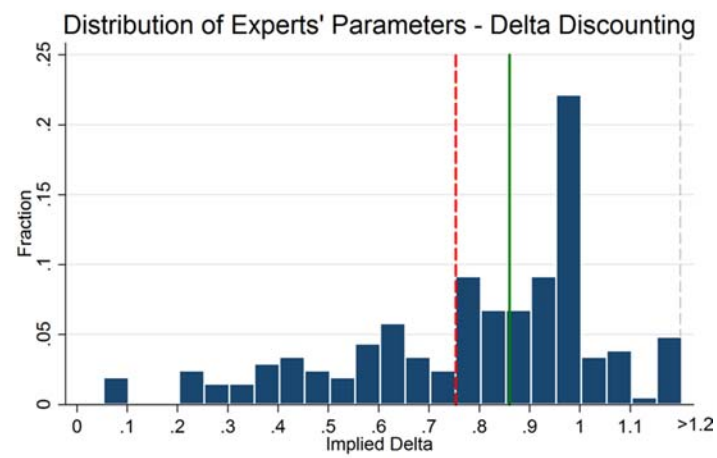


Red dashed line indicates the parameter value implied by the actual mTurker mean effort.
Green solid line indicates the median parameter value of the experts.
Parameter values can be interpreted as an intrinsic value
corresponding to paying an additional piece rate of x per 100 effort units.
Observations are truncated at 0 and 0.5.

Online Appendix Figures 9c-d. Estimate of Time Preference Parameters

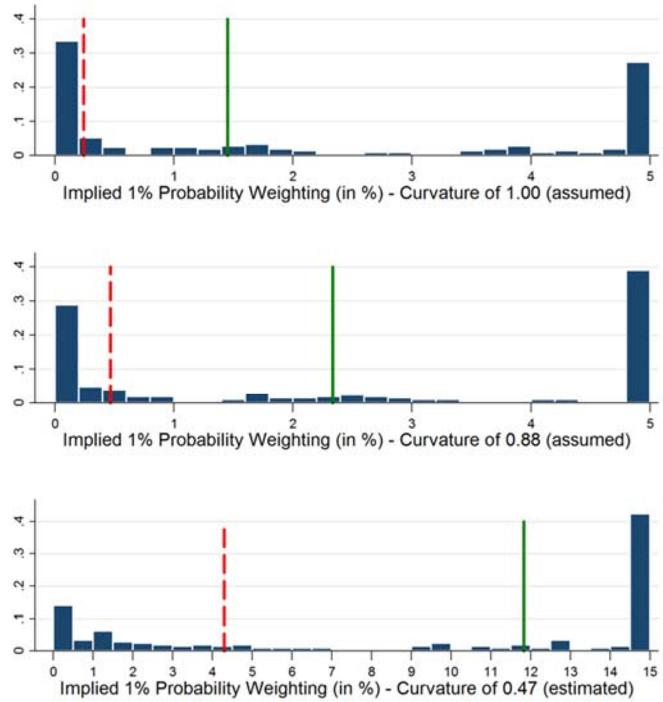


Red dashed line indicates the parameter value implied by the actual mTurker mean effort.
Green solid line indicates the median parameter value of the experts.
Observations are truncated at 1.2.

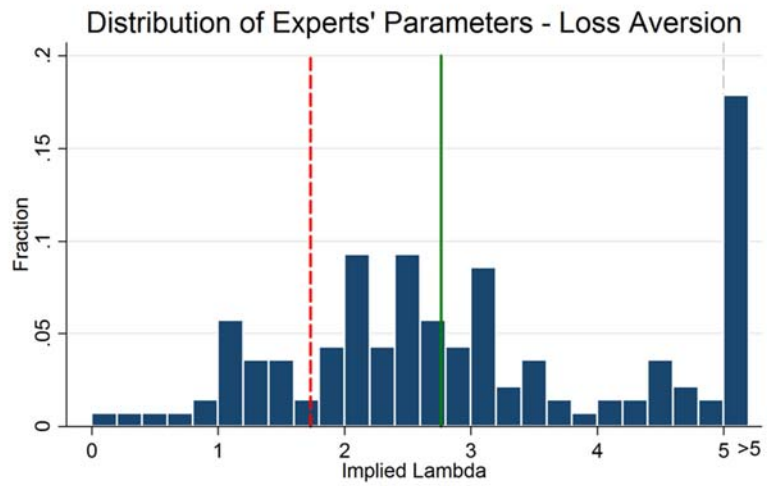


Red dashed line indicates the parameter value implied by the actual mTurker mean effort.
Green solid line indicates the median parameter value of the experts.
Observations are truncated at 1.2

Online Appendix Figures 9e-f. Estimate of Reference-Dependence Parameters



Implied probability weighting under three different specifications of the curvature over the piece rate (1, 0.88, 0.47). The red line indicates the actual probability weighting from the Mturkers' effort, the green line denotes the median forecast of the experts. Observations are truncated at 5, 5, and 15, respectively.



Red dashed line indicates the parameter value implied by the actual mTurker mean effort.
Green solid line indicates the median parameter value of the experts.
Observations are truncated at 5. Observations with a difference of less than 10 in the two gain-treatments and negative parameter values are dropped (i.e., 58 and 7 observations)

Notes: Online Appendix Figures 9a-f present the distribution of the estimates of the behavioral parameters from the relevant treatments (see Table 5). We use a minimum-distance estimator to estimate a model of costly effort with a power cost of effort function using the average effort in the three benchmark treatments for Online Appendix Figures 9a-d. The resulting parameter estimates are in Column (1), Panel A of Table 5. For Online Appendix Figure 9e we use a non-linear least squares estimate with an exponential cost function as in Table 6, Columns 4-6. Online Appendix Figure 9f is based on an approximate solution (see text). We use these estimated parameters and the observed effort in the relevant treatments to back out the implied structural estimate for a behavioral parameter from the relevant treatment (plotted as the red vertical line). Similarly, for each expert *i* we back out the expected behavioral parameter implied by the forecast which expert *i* makes for a particular treatment; the implied structural parameters are plotted in the figures, with the green line denotes the median parameter. See also the results in Panel B of Table 5. Online Appendix Figures 9a-b plot the implied altruism and warm glow parameters from the charitable giving treatments. Online Appendix Figure 9c-d plot the implied *beta* and *delta* from the time preference treatments. Online Appendix Figures 9e-f plot the implied probability weight (corresponding to a .01 probability) and loss aversion from the reference dependence treatments.

Online Appendix Table 1. Summary Statistics, Mturk Sample

	Mean	US Census
	(1)	(2)
Button Presses	1936	
Time to complete survey (minutes)	12.90	
US IP Address Location	0.85	
India IP Address Location	0.12	
Female	0.54	0.52
Education		
High School or Less	0.09	0.44
Some College	0.36	0.28
Bachelor's Degree or more	0.55	0.28
Age		
18-24 years old	0.21	0.13
25-30 years old	0.30	0.10
31-40 years old	0.27	0.17
41-50 years old	0.12	0.18
51-64 years old	0.08	0.25
Older than 65	0.01	0.17
Observations	9861	

Notes: Column (1) of Online Appendix Table 1 lists summary statistics for the final sample of Amazon Turk survey participants (after screening out ineligible subjects). Column (2) lists, where available, comparable demographic information from the US Census.

Online Appendix Table 2. Meta-Analysis of Findings in Literature, Individual Papers, Panel A

Category	Comparison	Paper	Outlet in economics	Google Scholar Citations	Subjects	Effort Task	Sample Size	Treatment	Effort in Treatment and Control, Mean(S.D.)	Treatment Effect in S.D., Cohen's d (S.e.)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Pay Enough or Don't Pay	Compare very-low-pay (1c per 1,000 points) to no piece rate	Gneezy and Rustichini (2000)	QJE	1800	Undergrads in Israel	Answer IQ questions	40 (T), 40 (C)	Pay just 10 cents NIS per correct answer (T) vs. no piece rate (C)	23.1 (14.7) (T), 28.4 (13.9) (C)	-0.372 (0.227)
		Gneezy and Rustichini (2000)	QJE	1800	High School Students in Israel	Fundraising	30 (T), 30 (C)	Pay just 1 percent of donations collected (T) vs. no commission (C)	154 (143) (T), 239 (166) (C)	-0.549 (0.268)
		Gneezy and Rey-Biel (2014)	JEEA	13	Consumers in the US	Survey Response	250 (T), 250 (C)	Pay $1 for completed survey (T) vs. no pay (C)	0.032 (0.176) (T), 0.076 (0.265) (C)	-0.196 (0.090)
		Charness, Cobo-Reyes and Sanchez (2016)	JEBO	2	Recruited from ORSEE	Staying to Enter More Data	29 (T), 30 (C)	2 cents piece rate during 2nd round (T) vs. no pay for 2nd round (C)	0.400 (0.490) (T), 0.241 (0.428) (C)	0.345 (0.264)
		Yang, Hsee and Urminsky (2014)	WP	3	University Research Lab	Look for Pennies Among Coins	58 (T), 58 (C)	Option to keep or donate pennies found (T) vs. no reward (C)	21.8 (10.1) (T), 17.0 (8.2) (C)	0.522 (0.192)
		Ashraf, Bandiera and Jack (2014)	JPubE	67	Hair stylists in Zambia	Selling Packs of Condoms	212 (T), 182 (C)	10 percent margin on their sales of condoms (T) vs. no incentives (C)	7.31 (13.98) (T), 6.93 (16.4) (C)	0.025 (0.101)
		Hossain and Li (2014)	MS	19	Students at HKUST	Data Entry	24 (T), 25 (C)	HK$0.50 piece rate (T) vs. no pay (C) (task described as work)	22.3 (4.1) (T), 24.2 (7.0) (C)	-0.331 (0.290)
		Hossain and Li (2014)	MS	19	Students at HKUST	Data Entry	24 (T), 24 (C)	HK$0.50 piece rate (T) vs. no pay (C) (task described as a favor for researchers)	21.5 (6.1) (T), 20.3 (5.1) (C)	0.223 (0.291)
Social Preferences: Charity	Compare low piece rate to charity (1c) to low piece rate to self (1c)	Imas (2014)	Journal of Public	37	University Students in the US	Squeeze a hand dynamometer	36 (C), 38 (T)	$0.05 piece rate to charity (T) vs. $0.05 piece rate by oneself (C) (in units of effort)	1.51 (0.87) (T), 1.14 (0.34) (C)	0.555 (0.242)
		Charness, Cobo-Reyes and Sanchez (2016)	JEBO	2	Recruited from ORSEE	Staying to Enter More Data	30 (C), 30 (T)	2 cents piece rate to charity (T) vs. 2 cents piece rate for oneself (C) (for 2nd round)	0.733 (0.442) (T), 0.400 (0.490) (C)	0.714 (0.275)
		Yang, Hsee and Urminsky (2014)	WP	3	University Research Lab	Look for Pennies Among	58 (C), 58 (T)	Pennies found to be donated (T) vs. option to keep or donate pennies found (C)	27.5 (11.4) (T), 21.8 (10.1) (C)	0.529 (0.192)
		Tonin and Vlassopoulos (2015)	MS	13	University Students in the UK	Data Entry	52 (C), 116 (C)	5p piece rate to charity (T) vs. 5p piece rate (C)	0.13 (0.31) (T), 0.08 (0.29) (C)	0.166 (0.167)
		Deehan et al (1997)	British Journal of	95	GPs in the UK	Survey Response	613 (C), 607 (T)	5 GBP to charity for survey completion (T) vs. 5 GBP for survey completion (C)	0.094 (0.292) (T), 0.171 (0.377) (C)	-0.230 (0.058)
	Compare high piece rate to charity (10c) to high piece rate to self (10c)	Imas (2014)	Journal of Public	37	University Students in the US	Squeeze a hand dynamometer	36 (T), 40 (C)	$2 piece rate to charity (T) vs. $2 piece rate by oneself (C) (in units of effort)	1.48 (1.03) (T), 1.74 (1.36) (C)	-0.217 (0.231)
		Charness, Cobo-Reyes and Sanchez (2016)	JEBO	2	Recruited from ORSEE	Staying to Enter More Data	30 (T), 30 (C)	8 cents piece rate to charity (T) vs. 8 cents piece rate for oneself (C) (for 2nd round)	0.80 (0.40) (T), 0.93 (0.25) (C)	-0.400 (0.263)
		Yang, Hsee and Urminsky (2014)	WP	3	University Research Lab	Look for Nickels Among Coins	55 (T), 55 (C)	Nickels found to be donated (T) vs. option to keep or donate nickels found (C)	24.1 (9.6) (T), 22.1 (9.7) (C)	0.207 (0.192)
		Tonin and Vlassopoulos (2015)	MS	13	University Students in the UK	Data Entry	52 (T), 100 (C)	10p piece rate to charity (T) vs. 10p piece rate (C)	0.12 (0.42) (T), 0.08 (0.29) (C)	0.105 (0.171)
		Deehan et al (1997)	British Journal of	95	GPs in the UK	Survey Response	598 (T), 578 (C)	10 GBP to charity for survey completion (T) vs. 10 GBP for survey completion (C)	0.100 (0.300) (T), 0.226 (0.418) (C)	-0.344 (0.059)
	Compare high piece rate to charity (10c) to low piece rate to charity (1c)	Imas (2014)	Journal of Public	37	University Students in the US	Squeeze a hand dynamometer	38 (C), 40 (T)	$2 piece rate to charity (T) vs. $0.05 piece rate to charity (C) (in units of effort)	1.48 (1.03) (T), 1.51 (0.87) (C)	-0.031 (0.227)
		Charness, Cobo-Reyes and Sanchez (2016)	JEBO	2	Recruited from ORSEE	Staying to Enter More Data	30 (C), 30 (T)	8 cents piece rate to charity (T) vs. 2 cents piece rate to charity (C) (for 2nd round)	0.800 (0.400) (T), 0.733 (0.442) (C)	0.158 (0.259)
		Yang, Hsee and Urminsky (2014)	WP	3	University Research Lab	Look for Pennies/Nickels	58 (C), 55 (T)	Nickels found to be donated (T) vs. pennies found to be donated (C)	24.1 (9.6) (T), 27.5 (11.4) (C)	-0.322 (0.191)
		Tonin and Vlassopoulos (2015)	MS	13	University Students in the UK	Data Entry	116 (C), 116 (T)	15p piece rate to charity (T) vs. 5p piece rate to charity (C)	0.14 (0.31) (T), 0.13 (0.31) (C)	0.033 (0.131)
		Deehan et al (1997)	British Journal of	95	GPs in the UK	Survey Response	607 (C), 578 (T)	10 GBP to charity for survey completion (T) vs. 5 GBP to charity for survey completion	0.100 (0.300) (T), 0.094 (0.292) (C)	0.022 (0.058)

Online Appendix Table 2. Meta-Analysis of Findings in Literature, Individual Papers, Panel B

Category	Comparison	Paper	Outlet in economics	Google Scholar Citations	Subjects	Effort Task	Sample Size	Treatment	Effort in Treatment and Control, Mean(S.D.)	Treatment Effect in S.D., Cohen's d (S.e.)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Social Preferences: Gift Exchange	Compare gift exchange (40c) to no piece rate	Gneezy and Rey-Biel (2014)	JEEA	13	Consumers in the US	Survey Response	250 (C), 1250 (T)	Gift ranging from $1 to $5 (T) vs. no pay (C)	0.193 (0.394) (T), 0.076 (0.265) (C)	0.311 (0.070)
		DellaVigna, List, Malmendier and Rao (2016)	WP	6	Temporary Workers from Craigslist	Stuff Envelopes	119 (C), 123 (T)	$14 pay (T) vs. $7 pay (C) (compared to $7 in the previous session)	37.1 (9.6) (T), 36.6 (8.5) (C)	0.050 (0.129)
		Gneezy and List (2006)	Econometrica	452	Undergraduates in the US	Data Entry	10 (C), 9 (T)	$20 hourly wage (T) vs. $12 hourly wage (C) (relative to the $12 advertised)	51.7 (15.5) (T), 40.7 (9.2) (C)	0.874 (0.506)
		Gneezy and List (2006)	Econometrica	452	Undergraduates in the US	Door-to-door Fundraising	10 (C), 13 (T)	$20 hourly wage (T) vs. $10 hourly wage (C) (relative to the $10 advertised)	10.0 (2.2) (T), 6.6 (2.3) (C)	1.51 (0.54)
		Bellemare and Shearer (2011)	International Economic Review	11	Planters working in British Columbia	Planting Trees	66 (C), 18 (T)	$80 gift and $0.20 piece rate (T) vs. $0.20 piece rate (C)	1153 (323) (T), 1063 (270) (C)	0.317 (0.268)
		Englmaier and Leider (2012)	WP	24	Temporary workers hired by HBS	Data Entry	14 (C), 15 (T)	$18 hourly wage (T) vs. $13 hourly wage (C) (relative to the $13 advertised)	23.5 (11.5) (T), 29.7 (16.1) (C)	-0.449 (0.382)
		Englmaier and Leider (2012)	WP	24	Temporary workers hired by HBS	Data Entry	15 (C), 15 (T)	$18 hourly wage (T) vs. $13 hourly wage (C) (relative to the $13 advertised). Subjects told that performance mattered to their	28.4 (8.4) (T), 24.4 (9.2) (C)	0.451 (0.375)
		Englmaier and Leider (2010)	WP	20	Recruited from CLER lab database at HBS	Solving puzzles on a computer	43 (C), 44 (T)	$20 hourly wage (T) vs. $10 hourly wage (C) (relative to the $10 advertised). Subjects told that performance mattered a little to their managers	202 (56) (T), 193 (48) (C)	0.180 (0.215)
		Englmaier and Leider (2010)	WP	20	Recruited from CLER lab database at HBS	Solving puzzles on a computer	53 (C), 52 (T)	$20 hourly wage (T) vs. $10 hourly wage (C) (relative to the $10 advertised). Subjects told that performance mattered a lot to their managers	204 (51) (T), 191 (67) (C)	0.222 (0.196)
		Kube, Marechal and Puppe (2012)	AER	207	Recruited from a German university	Data Entry	35 (C), 34 (T)	12 euro hourly wage and fixed payment of 7 euro (T) vs. 12 euro hourly wage (C) (relative to the 12 euro hourly wage	8742 (2605) (T), 8312 (1930) (C)	0.188 (0.242)
		Kube, Marechal and Puppe (2013)	JEEA	103	Recruited from a German university	Data Entry	25 (C), 22 (T)	20 euro hourly wage (T) vs. 15 euro hourly wage (C) (relative to the 15 euro advertised)	219 (135) (T), 219 (144) (C)	-0.006 (0.292)
		Esteves-Sorenson (2016)	WP	5	Students from 2 universities in the US	Data Entry	131 (C), 318 (T)	Raise for shift 1 and for some a subset, a raise for shift 3 (T) vs. no raise (C) (base hourly rate of $12)	17292 (6239) (T), 17591 (6917) (C)	-0.046 (0.104)
		Cohn, Fehr and Goette (2015)	MS	47	Workers in a Zurich publishing company	Distribute newspapers in public	178 (C), 181 (T)	Unexpected 27 CHF hourly rate for the shift (T) vs. 22 CHF hourly rate (C) (prior expectation was 22 CHF)	5.36 (0.40) (T), 5.35 (0.39) (C)	0.027 (0.106)
		Gilchrist, Luca and Malhotra (2016)	MS	10	Recruited from upwork.com	Enter CAPTCHAs	110 (C), 58 (T)	Unexpected net hourly wage of $4 (T) vs. net hourly wage of $3 (C) (subjects had requested wages between $2 and $3)	938 (420) (T), 792 (418) (C)	0.350 (0.165)

Online Appendix Table 2. Meta-Analysis of Findings in Literature, Individual Papers, Panel C

Category	Comparison	Paper	Outlet in economics	Google Scholar Citations	Subjects	Effort Task	Sample Size	Treatment	Effort in Treatment and Control, Mean(S.D.)	Treatment Effect in S.D., Cohen's d (S.e.)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Social Comparisons	Compare Cialdini-type comparison to no piece rate	Beshears et al. (2015)	Journal of Finance	114	Low-savings employees in the US	Enroll in savings plan with QE and 0% default	343 (C), 696 (T)	Info on savings and peer savings (T) vs. info on savings (C)	0.060 (0.236) (T), 0.099 (0.299) (C)	-0.150 (0.066)
		Beshears et al. (2015)	Journal of Finance	114	Low-savings employees in the US	Enroll in savings plan with QE and 6% default	136 (C), 264 (T)	Info on savings and peer savings (T) vs. info on savings (C)	0.027 (0.162) (T), 0.007 (0.083) (C)	0.142 (0.106)
		Beshears et al. (2015)	Journal of Finance	114	Low-savings employees in the US	Contribution rate to plan with EE and 0% default	235 (C), 511 (T)	Info on savings and peer savings (T) vs. info on savings (C)	0.106 (0.308) (T), 0.106 (0.308) (C)	0.000 (0.079)
		Beshears et al. (2015)	Journal of Finance	114	Low-savings employees in the US	Contribution rate to plan with EE and 6% default	931 (C), 1827 (T)	Info on savings and peer savings (T) vs. info on savings (C)	0.083 (0.276) (T), 0.082 (0.274) (C)	0.004 0.040)
		Bhargava and Manoli (2015)	AER	41	US tax filers who did not initially take up EITC	Take up EITC	20395 (C), 1753 (T)	Notice of eligibility and that similar peers are claiming (T) vs. notice of eligibility (C)	0.19 (0.39) (T), 0.23 (0.42) (C)	-0.096 (0.025)
		Cai et al. (2009)	AER	199	Restaurant visitors in China	Purchase a top dish	1772 (C), 2182 (T)	Plaque displaying 5 top dishes (T) vs. nothing displayed on diners' tables (C)	0.183 (0.387) (T), 0.162 (0.368) (C)	0.055 (0.032)
		Coffman et al. (2017)	AEJ: Applied	4	Applicants to Teach for America	Accept job offer	3337 (C), 3348 (T)	Admission letter with line on social norm (T) vs. standard admission letter (C)	0.790 (0.407) (T), 0.773 (0.419) (C)	0.041 (0.024)
		Fellner et al. (2013)	JEEA	136	Potential evaders of TV license fines in Austria	Respond to Mail Notice	7984 (C), 7998 (T)	Warning letter and social information (T) vs warning letter (C)	0.407 (0.491) (T), 0.431 (0.495) (C)	-0.048 (0.016)
		Fellner et al. (2013)	JEEA	136	Potential evaders of TV license fines in Austria	Respond to Mail Notice	7821 (C), 8101 (T)	Warning letter, threat and social information (T) vs warning letter and threat (C)	0.428 (0.495) (T), 0.450 (0.498) (C)	-0.045 (0.016)
		Frey and Meier (2004)	AER	695	Students at the University of Zurich	Donate to Charitable Fund	500 (C), 1000 (T)	Contribution form and info about high social norm (T) vs. contribution form	0.770 (0.421) (T), 0.729 (0.444) (C)	0.096 (0.055)
		Goldstein et al. (2008)	Journal of Consumer Research	1227	Guests at a well-known hotel chain in the US	Reuse towel	216 (C), 217 (T)	Social norm message (T) vs. typical request to reuse towesl (C)	0.441 (0.497) (T), 0.351 (0.477) (C)	0.185 (0.097)
		Goldstein et al. (2008)	Journal of Consumer Research	1227	Guests at a well-known hotel chain in the US	Reuse towel	319 (C), 1276 (T)	Social norm message (T) vs. typical request to reuse towesl (C)	0.445 (0.497) (T), 0.372 (0.483) (C)	0.148 (0.063)
		Hallsworth et al. (2014)	NBER WP	89	Originally non-compliant UK taxpayers	Comply with tax payment	16912 (C), 50735 (T)	Standard letter and one of 3 norm treatments (T) vs. standard letter	0.354 (0.478) (T), 0.336 (0.472) (C)	0.037 (0.009)
		Hallsworth et al. (2014)	NBER WP	89	Originally non-compliant UK taxpayers	Comply with tax payment	8538 (C), 93918 (T)	Standard letter and one of 11 norm treatments (T) vs. standard letter	0.365 (0.481) (T), 0.336 (0.472) (C)	0.061 (0.011)
		Krupka and Weber (2009)	Journal of Economic Psychology	128	Students at Carnegie Mellon and University of Pittsburgh	Prosocial choice in dicatator game	38 (C), 120 (T)	Information on others' behavior (T) vs. no information	0.54 (0.50) (T), 0.34 (0.47) (C)	0.406 (0.189)

Online Appendix Table 2. Meta-Analysis of Findings in Literature, Individual Papers, Panel D

Category	Comparison	Paper	Outlet in economics	Google Scholar Citations	Subjects	Effort Task	Sample Size	Treatment	Effort in Treatment and Control, Mean(S.D.)	Treatment Effect in S.D., Cohen's d (S.e.)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Probability Weighting	Compare probabilistic piece rate (1% of $1) to deterministic piece rate with expected value (1c)	Halpern et al. (2011)	Health Services Research	24	Resident Physicians in a US Database	Survey Response	400 (C), 358 (T)	0.4% chance of winning US$2500 (T) vs. fixed payment of US$10 (C) for response	0.511 (0.500) (T), 0.558 (0.497) (C)	-0.093 (0.073)
		Thirumurthy et al. (2016)	J. Acquired Immune Deficiency Syndromes	2	Men aged 21 to 39 years old in Kenya	Uptake of Circumcision	308 (C), 302 (T)	Mixed lottery with expected retail value of US$12.50 (T) vs. food voucher worth US$12.50 (C)	0.033 (0.179) (T), 0.084 (0.278) (C)	-0.219 (0.081)
		Diamond and Loewy (1991)	J. Applied Social Psych.	53	Undergraduates in State University	Recycling	113 (C), 78 (T)	5% chance of winning $5 and 1% chance of winning $25 (T) vs. $0.50 voucher for campus store (C)	0.308 (0.462) (T), 0.212 (0.409) (C)	0.221 (0.148)
		Dolan and Rudisill (2014)	Social Science & Medicine	2	16 to 24 year olds in England	Return Test Kit via Mail	549 (C), 247 (T)	10% chance of a 50 GBP Tesco voucher (T) vs. 5 GBP Tesco voucher (C)	0.706 (0.455) (T), 0.732 (0.443) (C)	-0.058 (0.077)
Ranking	Compare expectation of rank to no piece rate	Kosfeld and Neckermann (2011)	AEJ Micro	167	Students in Zurich	Online Search and Data Entry	83 (C), 67 (T)	Fixed pay and possibility of award based vaguely on award (T) vs. fixed pay (C)	0.253 (0.090) (T), 0.226 (0.059) (C)	0.363 (0.167)
		Barankay (2012)	WP	59	Furniture Salespeople in North America	Sales Performance	439 (C), 439 (T)	Rank feedback expected (T) vs. no rank feedback expected (C)	8.58 (1.02) (T), 8.78 (0.95) (C)	-0.204 (0.068)
		Ashraf, Bandiera and Lee (2014)	JEBO	39	Civil Service Cadre Trainees in Zambia	Exam Scores	61 (C), 247 (T)	Individual and rank feedback expected (T) vs. only individual feedback expected (C)	-0.188 (1.698) (T), 0.000 (1.000) (C)	-0.119 (0.143)
		Blanes i Vidal and Nossol (2011)	MS	110	Warehouse Workers in Germany	Warehouse Tasks Completed	57 (C), 59 (T)	Rank feedback expected (T) vs. no rank feedback expected (C)	5.01 (0.12) (T), 4.96 (0.12) (C)	0.387 (0.189)
		Gill, Kissova, Lee and Prowse (2016)	WP	6	Students at the University of Oxford	Verbal and Numerical Tasks	51 (C), 255 (T)	Individual and rank feedback expected (T) vs. only individual feedback expected (C)	74.1 (19.6) (T), 67.4 (19.1) (C)	0.343 (0.155)
Task Significance	Compare task significance to no piece rate	Grant (2008)	Journal of Applied Psychology	452	Callers at fundraising organization	Solicit donations	11 (C), 12 (T)	Read stories about beneficiaries (T) vs. fill in surveys (C)	23.0 (11.4) (T), 10.1 (4.6) (C)	1.46 (0.53)
		Grant (2008)	Journal of Applied Psychology	452	New callers at fundraising organization	Solicit donations	17 (C), 17 (T)	Read stories about beneficiaries (T) vs. fill in surveys (C)	27.9 (13.7) (T), 10.1 (4.6) (C)	0.695 (0.364)
		Grant et al. (2007)	OB and Human Decision P.	255	Callers at fundraising organization	Solicit donations	10 (C), 12 (T)	Read letter by beneficiary and discussed between themselves (T) vs. no contact (C)	147 (58) (T), 179 (57) (C)	-0.558 (0.446)
		Grant et al. (2007)	OB and Human Decision P.	255	Callers at fundraising organization	Solicit donations	10 (C), 17 (T)	Talked to beneficiary (T) vs. no contact (C)	261 (135) (T), 179 (57) (C)	0.722 (0.424)
		Chandler and Kapelner (2013)	JEBO	175	MTurk workers from US and India	Image labelling	798 (C), 845 (T)	Subjects told that they were labelling tumor cells to assist medical research (T) vs. no such information (C)	0.806 (0.395) (T), 0.762 (0.426) (C)	0.107 (0.049)
		Grant (2012)	Ac. of Management Journal	212	New employees at a call center in the US Midwest	Sales of educational and marketing	26 (C), 45 (T)	Visit by director and/or benficiary (T) vs. no visit (C)	180 (87) (T), 46 (39) (C)	1.82 (0.33)
		Ariely, Kamenica and Prelec (2008)	JEBO	116	MIT students	Matching letters on sheets	35 (C), 34 (T)	Subjects told to put their names on their sheets (T) vs. subjects told not to do so (C)	9.03 (2.41) (T), 6.77 (2.50) (C)	0.921 (0.266)

Notes: The Table lists the papers in the meta-analysis of related treatments. We require: (i) a laboratory or field experiment (or natural experiment); (ii) a treatment comparison that matches the one in our study; (iii) an outcome variable about (broadly conceived) effort, such as responding to a survey. For each treatment, we specify a comparison of treatments.

Online Appendix Table 2. Meta-Analysis of Findings in Literature, Notes, Panel E

Treatment	Paper	Notes
Paying Too Little versus No Pay	Gneezy and Rustichini (2000)	We computed Cohen's d for 2 separate experiments based on values reported in the text and in tables 1 and 4.
	Gneezy and Rey-Biel (2014)	We computed Cohen's d based on values reported in the text and in table A.1.
	Gneezy and Rey-Biel (2014)	We consider the $1 pay treatment as very low pay, given that $1 pay for a 15 minute survey was low pay for most typical US consumers.
	Charness, Cobo-Reyes and Sanchez (2016)	Subjects in the first stage enter on average 120 entries in one hour, so the 2 cents piece rate translates into $2.40 per hour pay for staying for the 2nd round. We decided that this pay was sufficiently low.
	Yang, Hsee and Urminsky (2014)	Participants in the "own piece rate group" also had an option to donate. The exact sample sizes for the treatment and control groups separately are not apparent based on the text, so we assumed they were equally sized.
	Ashraf, Bandiera and Jack (2014)	The financial incentive is equivalent to about USD 0.01 per pack of condom sold, where the mean number of packs sold over the entire study period (one year) is about 9. So, we categorize this as a very low financial incentive.
	Hossain and Li (2014)	The two treatment-control comparisons from this paper differ in that in one comparison, the task was described to both control and treatment groups purely as work (which the authors call the work frame), whereas in the second comparison, it was described to both groups as a favor to researchers (which the authors call the social frame).
Charity	Tonin and Vlassopoulos (2015)	Statistics were calculated based on values reported in table 4, as well as summary data kindly provided to us by the authors.
	Deehan et al (1997)	This was a selected sample of doctors (GPs) in that the GPs in both the control and treatment arms that we define had not responded to the initial 2 waves of the survey (for which response was not incentivized)
Ranking versus No Pay	Kosfeld and Neckermann (2011)	There were two main measures of effort -- number of communities the subjects entered per minute, and the number of points the subjects scored per minute. We use the former measure because it was easier to interpret.
	Barankay (2012)	All individuals at this firm used to have rank feedback, and the experimental intervention removed this feedback for some. So, this is slightly different from other papers in this category where the "default" is typically no rank feedback.
	Ashraf, Bandiera and Lee (2014)	We defined our treatment group as treatments 1 to 4 pooled (since they all included various elements of ranking). We also only focused on the results from the first exam, since the subjects received rank feedback subsequently. Finally, we derived the treatment group standard deviation using the regression in column (3) of Table 2 which control for subject characteristics (since a regression without these controls was not reported). The outcome variable was normalized by the mean and standard deviation in the control group.
	Blanes i Vidal and Nossol (2011)	Notice that this is a quasi-field experiment, with a time series switch over time. We used unweighted averages of individuals' daily productivity during the period before the firm announced to workers that they will be receiving information about their individual rank, and during the period after the firm announced to the workers but before workers actually started receiving rank feedback, as our "control and treatment groups" respectively. The data required to compute these values was kindly provided to us by the authors.
	Gill, Kissova, Lee and Prowse (2016)	We only used data from the first round, since subjects subsequently started receiving rank feedback. The data for the first round was kindly provided to us by the authors.
Task Significance	Grant (2012)	We pool the treatment arms for the visit and speech by the director and/or the beneficiary since these all treatments all conveyed to subjects (in different ways) the significance of their work.
	Ariely, Kamenica and Prelec (2008)	We did not include the treatment where subjects' sheets were shredded since this was a form of "negative" task significance that is quite different from the other task significance treatments.
Cialdini Comparison	Beshears et al. (2015)	The abbreviations QE and EE in treatment summary tables stand for Quick Enrollment and Easy Escalation respectively.
	Frey and Meier (2004)	The sample sizes we listed are in fact upper bounds, since there was some sample attrition due to students not reenrolling (the authors only reported the numbers before attrition).
	Goldstein et al. (2008)	Slightly different language was used in the two control/treatment comparisons we extracted from this paper.
	Hallsworth et al. (2014)	We used two control/treatment comparisons from this paper. In the first, we combined the results for the 3 norm conditions, with the sample size based on the total sample split 6 ways equally (3 parts social norms; 1 part control), and taking the average effect across social norms from table 4. For the second, we combined the results for the 11 norm conditions, with the sample size based on the total sample split 13 ways, and taking the average effect across all social norms in Table 7.
Probability weighting	Thirumurthy et al. (2016)	The mixed lottery consisted of a 5% chance of winning a bicle or smartphone worth US $120, 10% chance of winning a standard mophile phone or pair of shoes worth US $45 and 85% chance of winning a food voucher worth US $2.50 (expected value of lottery = $12.50), conditional on undergoing circumcision within 3 months. A potential concern with the comparability of expected values in the control versus treatment groups is that subjects' willingness to pay for some of these items may be lower than the items' retail prices.
	Diamond and Loewy (1991)	The randomization in this paper occurred at the dormitory level. We use the data for the earlier December collection period for our analysis.
Gift Exchange vs. No Pay	Englmaier and Leider (2012)	We coded two treatment/control comparisons for this paper. While both compared the effect of a monetary gift on performance, in one case subjects (in both the control and treatment groups) were told that their managers will get a substantial "completion bonus" if enough work gets done. We used number of characters of data entered per minute as the outcome variable. The results obtained using instead the accuracy-corrected rate as dependent variable were qualitatively similar.
	Englmaier and Leider (2010)	We coded two treatment/control comparisons for this paper. While both compared the effect of a monetary gift on performance, in one case subjects (in both the control and treatment groups) were told that their managers' payoff depended to a large extent on their performance whereas in the other case subjects were told that their managers' payoff depended on their performance only to a small extent.
	Kube, Marechal and Puppe (2012)	This meta-analysis includes only the monetary gift arms, not the in-kind gifts, which are not comparable to our treatments.
	Kube, Marechal and Puppe (2013)	We take the sample sizes and means for the control and treatment groups on pages 858 and 859. Since the standard deviations were not reported in the table, we approximated them using the standard error for the constant from the regression in column 1 of Table 2.
	Esteves-Sorenson (2016)	Some students in the 67% raise group were told 1 week in advance that they were getting the raise, whereas some got the news immediately before starting the task. Similarly to the authors, we pool the "67% raise before shift 1" group with "the 50% raise before shift 1 (then possibly raised again to 100% before shift 3)" group.
	Cohn, Fehr and Goette (2015)	We computed the means in the control and treatment groups using data that the authors made available online, using log hourly copies as our outcome variable and dropping observations with missing values of this variable. We use the number of workers who experienced a control/treatment shift as the number of observations in the control/treatment groups.

Online Appendix Table 3. Meta-Analysis of Probability Weighting Estimates in Literature

Paper	Outlet in economics	Google Scholar Citations	Setting	Type of Probability Weighting Function	Parameter Estimate	Implied Probability Weight for 1% Probability	Implied Probability Weight for 50% Probability
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Kahneman and Tversky (1992)	J Risk and Uncertainty	9502	Lottery Choice	Kahneman-Tversky	0.61	0.055	0.421
Gonzalez and Wu (1999)	Cognitive Psychology	798	Lottery Choice	Linear-in-log-odds	-	0.093 (0.003)	0.435 (0.010)
Camerer and Ho (1994)	JRU	662	Lottery Choice	Kahneman-Tversky	0.56	0.067	0.393
Gonzalez and Wu (1996)	MS	714	Lottery Choice	Kahneman-Tversky	0.71	0.036 (0.002)	0.461 (0.010)
Harrison, List and Towe (2007)	EMA	166	Lottery Choice	Kahneman-Tversky	0.83	0.022	0.488
Kilka and Weber (2001)	MS	215	Stock Forecasts	Linear-in-log-odds	-	0.181 (0.013)	0.481 (0.002)
Abdellaoui (2000)	MS	687	Lottery Choice	Linear-in-log-odds	0.6	0.040 (0.001)	0.394 (0.007)
Tversky and Fox (1995)	Psychological Review	904	NBA/NFL/Weather Forecasts	Linear-in-log-odds	-	0.031	0.435
Donkers, Melenberg and van Soest (2001)	JRU	335	Lottery Choice	Prelec	0.435	0.143 (0.011)	0.426 (0.001)
Harrison, Humphrey and Verschoor (2010)	Economic Journal	143	Lottery Choice	Kahneman-Tversky	1.384	0.002 (0.000)	0.464 (0.002)
Bruhin, Fehr-Duda and Epper (2010)	Econometrica	223	Lottery Choice	Linear-in-log-odds	-	0.141 (0.003)	0.481 (0.001)
de Brauw and Eozenou (2014)	J Dev. Economics	32	Crop Choice	Kahneman-Tversky	1.37	0.002 (0.000)	0.467 (0.001)
Liu (2013)	REStat	135	Lottery Choice	Prelec	0.69	0.057 (0.014)	0.460 (0.004)
Tanaka, Camerer and Nguyen (2010)	AER	472	Lottery Choice	Prelec	0.74	0.045	0.467
Barseghyan, Molinari, O'Donoghue and Teitelbaum (2013)	AER	112	Insurance Deductible Choice	Semi-nonparametric	-	0.07	-
Snowberg and Wolfers (2011)	JPE	180	Horse Race Data	Prelec	0.928	0.020	0.491
Aruoba and Kearny (2011)	Working paper	5	State Lotteries	Prelec	0.89	0.020	0.486
Liger and Levy (2009)	JEBO	35	Financial Markets	Kahneman-Tversky	0.622	0.053 (0.001)	0.426 (0.003)
Average Probability Weight from Meta-Analysis						**π(0.01) = 0.060**	**π(0.50) = 0.452**
Implied Effort in Probabilistic Pay Treatments (Assuming Linear Value Function)						**2,142 points (1% of $1)**	**2,023 points (50% of 2c)**
Implied Effort in Probabilistic Pay Treatments (Assuming Curvature of 0.88 as in TK)						**2,117 points (1% of $1)**	**2,016 points (50% of 2c)**
Implied Effort in Probabilistic Pay Treatments (Assuming Curvature of 0.7)						**2,065 points (1% of $1)**	**2,002 points (50% of 2c)**

Notes: The Table lists papers providing an estimate of the probability weighting function, with the paper and journal (Columns 1 and 2), the Google Scholar citations (Column 3), the setting and type of probability weighting function used (Columns 4 and 5), and the estimated parameter for the probability weighting function, when available (Column 6). The key columns are Column 7 and 8, which report the implied probability weight for a 1 % probability and a 50% probability, given the estimated weighting function in the study. The standard errors, when available, are computed with the delta method. At the bottom of the table we report the parameter for the meta-analysis, equal-weighting across the studies. We also report the implied average effort (point) in the 1% treatment and 50% treatment, assuming different degrees of curvature in the utility function. For the case of no curvature, we take the benchmark estimates of the parameters in Table 5, Column 1, while for the case of curvature we re-estimate the model with minimum-distance on the 3 benchmark moments with the assumed degree of curvature.

Online Appendix Table 4. Estimates of Behavioral Parameters, Robustness

Cost of Effort Specification:	Exponential Cost of Effort			
Estimation Method:	Non-linear Least Squares Estimator on Individual Effort			
Assumption:	*Low Cost Function Curvature*	*High Cost Function Curvature*	*Concave Value Function*	*Continuous Points*
	(1)	(2)	(3)	(4)
Panel A. Estimate of Model on Effort in 3 Benchmark Treatments				
Curvature γ of Cost of Effort Function	0.010 (assumed)	0.020 (assumed)	0.0138 (0.003)	0.0159 (0.0040)
Level *k* of Cost of Effort Function	2.41E-11 (4.46E-12)	1.80E-20 (6.61E-21)	1.34E-14 (9.78E-14)	1.05E-16 (8.92E-16)
Intrinsic Motivation *s* (cent per point)	9.86E-05 (3.59E-05)	2.98E-07 (2.16E-07)	1.67E-05 (3.49E-05)	3.13E-06 (7.63E-06)
Curvature of Utility Over Piece Rate	1 (assumed)	1 (assumed)	0.88 (assumed)	1 (assumed)
N	1664	1664	1664	1664
Implied Effort, 4-cent Treatment (Actual Effort 2,132)	2123	2112	2115	2117
Implied Effort, Low-pay Treatment (Actual Effort 1,883)	1754	1928	1880	1884

Panel B. Estimates of Social Preferences and Time Preferences	**Estimate from Mturk (95% c.i.)**	**Median Forecast (25th, 75th ptile)**	**Estimate from Mturk (95% c.i.)**	**Median Forecast (25th, 75th ptile)**	**Estimate from Mturk (95% c.i.)**	**Median Forecast (25th, 75th ptile)**	**Estimate from Mturk (95% c.i.)**	**Median Forecast (25th, 75th ptile)**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Social Preferences Parameters								
Pure Altruism Coefficient α	0.007	0.1	0.002	0.053	0.007	0.12	0.003	0.067
	(-0.032,0.046)	(0.01,0.34)	(-0.01,0.014)	(0.001,0.69)	(-0.033,0.047)	(0.006,0.67)	(-0.017,0.024)	(0.002,0.53)
Warm Glow Coefficient *a*	0.432	0.4	0.060	0.003	0.313	0.19	0.14	0.031
	(0.125,0.738)	(0.042,1.4)	(-0.025,0.145)	(-0.006,0.42)	(-0.216,0.842)	(0.004,1.5)	(-0.13,0.41)	(-0.0002,0.7)
Gift Exchange Δ*s*	3.03E-04	3.00E-04	3.0E-06	4.7E-06	8.5E-05	1.0E-04	2.1E-05	2.7E-05
(cent per point)	(-6.7E-5,7E-4)	(1E-4,1.5E-3)	(-2E-6,8E-6)	(6E-7,8E-5)	(-2E-4,4E-4)	(2E-5,8E-4)	(-6E-5,1E-4)	(4E-6,2.7E-4)
Time Preference Parameters								
Present Bias β	1.74	1.3	0.95	0.54	1.21	1.4	1.15	0.76
	(-0.48,3.97)	(0.7,1.7)	(-1.46,3.36)	(0.16,0.93)	(-0.971,3.38)	(0.6,2.1)	(-1.24,3.54)	(0.28,1.2)
(Weekly) Discount Factor δ	0.83	0.91	0.70	0.82	0.78	0.87	0.76	0.85
	(0.51,1.16)	(0.75,1)	(0.16,1.24)	(0.58,1)	(0.35,1.21)	(0.68,1)	(0.28,1.25)	(0.65,1)

Notes: This table reports the results of four robustness checks, each estimated using a non-linear least squares estimator with an exponential cost of effort function. The specification regresses the effort of the individual MTurker (rounded to the nearest 100 points) with the specification discussed in Section 6. The specification in Panel A include only the 3 benchmark treatments, while the specifications in Panel B include also the charitable giving, gift exchange, and time-delay treatments. For each specification, the first Column in Panel B presents the parameter estimates from the MTurker effort, while the second column presents the implied parameter value for the expert forecast at the median, the 25th percentile and the 75th percentile of the expert distribution. The first two robustness checks examine the impact of mis-specifications in the cost of effort function by forcing the curvature parameter to be fixed at a low value (Column 1) or a high value (Cloumn 2). The second robustness check involves estimates which assume a concave value function, as opposed to linear utility, taking the Tversky and Kahneman 0.88 curvature. Column 4 is like the benchmark, except that, instead of using the points rounded to 100, it uses the continuous points, assuming (for simplicity) that the incentives are distributed continuously.